

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 09:49:29 ; Search time 4371 Seconds

(Without alignments)
10239.106 Million cell updates/sec

Title: US-09-940-227-15

Perfect score: 1094
Sequence: 1 gtccacgggggactgttacc.....tggtttttaaaaaaaaaa 1094

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrl:*
38: em_gy:*
39: em_hlg_hum:*
40: em_hlg_mus:*
41: em_hlg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1094	100.0	1094	6	AX472966	AX472966 Sequence
2	799	73.0	4520	9	AF151978	AF151978 Homo sapi
3	563	51.5	3007	10	AF320226	AF320226 Mus muscu
4	562.2	51.4	2337	10	AF161714	AF161714 Mus muscu
5	505.8	46.2	1923	10	AB033285	AB033285 Mus muscu
6	392	35.8	9682	9	HS452H17	Z96810 Human DNA s
7	334.4	30.6	178704	2	BX119904	BX119904 Human DNA
8	305.8	28.0	151973	2	AC133215	AC133215 Homo sapi
9	305.8	28.0	174655	9	AC023347	AC023347 Homo sapi
10	246	22.5	158527	2	AC027640	AC027640 Homo sapi
11	246	22.5	178704	9	BX119904	BX119904 Human DNA
12	171.4	15.7	141025	10	AL807250	AL807250 Mouse DNA
13	171.2	15.6	195026	2	AC112474	AC112474 Rattus no
14	171.2	15.6	216341	2	AC134727	AC134727 Rattus no
15	171.2	15.6	323793	2	AC108660	AC108660 Rattus no
16	164	14.1	2544	3	AK112740	AK112740 Clona Int
17	154.2	14.1	6982	10	AF411042	AF411042 Mus muscu
18	154.2	14.1	7081	10	AY147186	AY147186 Mus muscu
19	149.4	13.7	2394	6	AR217626	AR217626 Sequence
20	149.4	13.7	2394	6	AR217627	AR217627 Sequence
21	149.4	13.7	2450	9	AF142501	AF142501 Homo sapi
22	147.8	13.5	2394	6	AR217628	AR217628 Sequence
23	147.8	13.5	2394	9	AF117999	AF117999 Homo sapi
24	147.8	13.5	2394	9	AF352733	AF352733 Homo sapi
25	147.8	13.5	2729	9	AF085412	AF085412 Homo sapi
26	147.8	13.5	2863	6	AR154269	AR154269 Sequence
27	147.8	13.5	2863	6	AR221682	AR221682 Sequence
28	147.8	13.5	2863	6	E38403	E38403 SC6 polypep
29	136.6	12.5	2817	6	AR048966	AR048966 Sequence
30	136.6	12.5	2817	6	AR080685	AR080685 Sequence
31	136.6	12.5	2817	10	RATGTX	L21672 Rat mRNA se
32	107.8	9.9	137095	9	AC135452	AC135452 Homo sapi
33	107.8	9.9	158527	2	AC027640	AC027640 Homo sapi
34	101.4	9.3	453	6	AR217611	AR217611 Sequence
35	101.4	9.3	453	6	AR217612	AR217612 Sequence
36	99.8	9.1	453	6	AR217610	AR217610 Sequence
37	99.8	9.1	453	6	AR217613	AR217613 Sequence
38	99.8	9.1	453	6	AR217614	AR217614 Sequence
39	99.8	9.1	453	6	AR217615	AR217615 Sequence
40	99.8	9.1	453	6	AR217616	AR217616 Sequence
41	99.8	9.1	453	6	AR217617	AR217617 Sequence
42	86.6	7.9	755	6	AR217602	AR217602 Sequence
43	86.6	7.9	755	6	AR217603	AR217603 Sequence
44	86.6	7.9	755	6	AR217604	AR217604 Sequence
45	86.6	7.9	755	6	AR217605	AR217605 Sequence

ALIGNMENTS

RESULT 1
AX472966
LOCUS AX472966 1094 bp DNA
DEFINITION Sequence 15 from Patent WO0218576.
ACCESSION AX472966
VERSION AX472966.1 GI:22207753
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Chen, S.Y., MacIua, R.A., Sun, Y. and Recipon, H.
AUTHORS
TITLE Compositions and methods relating to lung specific genes
JOURNAL Patent: WO 0218576-A 15 07-MAR-2002;

FEATURES		Diadexus, Inc. (US)	
Source	Location/Qualifiers		
	1. 1094		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	338 a 172 c 236 g 348 t		
ORIGIN			
Query Match	100.0% Score 1094; DB 6; Length 1094;		
Best Local	Similarity 100.0%; Pred. No. 2,5e-225;		
Matches 1094; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
1	GTTCACAGGGGACTGTTACCTTACAGTTGTTATCGATGAAAAATCATATTAAGCAGACTA	60	
1	GTTCACAGGGGACTGTTACCTTACAGTTGTTATCGATGAAAAATCATATTAAGCAGACTA	60	
61	AAACAGAAATPAGGGCTGCATTTGATTTCTATCGAACCAGAGTGTGACTCCAAAAGA	120	
61	AAACAGAAATPAGGGCTGCATTTGATTTCTATCGAACCAGAGTGTGACTCCAAAAGA	120	
121	AAAAACCCGTGGCTTTTAAAGCAAGAAACAGAGGCTCAGAGTCTATTTTCAACATATA	180	
121	AAAAACCCGTGGCTTTTAAAGCAAGAAACAGAGGCTCAGAGTCTATTTTCAACATATA	180	
181	TTTGTAATTTGATTTAATACATACATGTAATTCGTATTAACGTAATTTGGCAATTAAGTTG	240	
181	TTTGTAATTTGATTTAATACATACATGTAATTCGTATTAACGTAATTTGGCAATTAAGTTG	240	
241	ACATATTAACATGTCAATTTCTCCCTCTGTGAAGAAACGATCAGACAAACATTTCAAGATT	300	
241	ACATATTAACATGTCAATTTCTCCCTCTGTGAAGAAACGATCAGACAAACATTTCAAGATT	300	
301	ATTGCCAAAGATGATGAAGAAATAGGGTCCCAATCTTTGGGCTGCTGGTGGTTT	360	
301	ATTGCCAAAGATGATGAAGAAATAGGGTCCCAATCTTTGGGCTGCTGGTGGTTT	360	
361	GTTCCTCCTTGGTCTCGTCTGTGTGACTCAGGCTGGAAATTTACTGGGTTCACTGTGATGA	420	
361	GTTCCTCCTTGGTCTCGTCTGTGTGACTCAGGCTGGAAATTTACTGGGTTCACTGTGATGA	420	
421	CCACTTCGTGTGCTGATGGGGCAATTTTAATTCAGCTATACCTGAGACTAGTTGATCAT	480	
421	CCACTTCGTGTGCTGATGGGGCAATTTTAATTCAGCTATACCTGAGACTAGTTGATCAT	480	
481	CTGATTTATGAGAGGACAGATTCATTTAGAGATACAGAAATGATGATTTGGACCAAGAG	540	
481	CTGATTTATGAGAGGACAGATTCATTTAGAGATACAGAAATGATGATTTGGACCAAGAG	540	
541	GTGATATTTCTGGCTATGTGTGAGACCTGCTGGTGTGTAATTAAGCCCTATCTTTGAT	600	
541	GTGATATTTCTGGCTATGTGTGAGACCTGCTGGTGTGTAATTAAGCCCTATCTTTGAT	600	
601	TGCAATATTTATCTGTGCTATTGGTGTGCAATTTCTATAGACCTAATTTATGGGCAATTCATA	660	
601	TGCAATATTTATCTGTGCTATTGGTGTGCAATTTCTATAGACCTAATTTATGGGCAATTCATA	660	
661	CCCTGACTGGGGAGTTGCTTTAGCGTGCGGTATGATGTTTCTGATTAATTTGGATTC	720	
661	CCCTGACTGGGGAGTTGCTTTAGCGTGCGGTATGATGTTTCTGATTAATTTGGATTC	720	
721	AATTATGGCTATCATTAATAATTAATTAAGGCTTAAGGAAACATCTTTCACAGCCCTTAAG	780	
721	AATTATGGCTATCATTAATAATTAATTAAGGCTTAAGGAAACATCTTTCACAGCCCTTAAG	780	
781	TTGCTCAGACACAGCTTCTAATCGGGGTCCATACCTGGAACAACATCTGTGGGAAAGATA	840	
781	TTGCTCAGACACAGCTTCTAATCGGGGTCCATACCTGGAACAACATCTGTGGGAAAGATA	840	
841	TAAAGCATGCTAGTCTCTAATAAAGAGGCTGGCATGAATTAACCTTACTTTAGTGGCAG	900	
841	TAAAGCATGCTAGTCTCTAATAAAGAGGCTGGCATGAATTAACCTTACTTTAGTGGCAG	900	

QY	901	CAGAAACCCGGAATGAGATCTCTATTGAAAAAAATATATGATGTATTAAGTGAATTTTTT	960
Db	901	CAGAAACCCGGAATGAGATCTCTATTGAAAAAAATATATGATGTATTAAGTGAATTTTTT	960
QY	961	TAGATATAGGGGAGACCCTTATTTATTTGTGTGTTAACTGATAGCAAAATGTACATCTAT	1020
Db	961	TAGATATAGGGGAGACCCTTATTTATTTGTGTGTTAACTGATAGCAAAATGTACATCTAT	1020
QY	1021	GTTTCATGATAGGGGATGATTTTTCCTCCATTAAACAGAAATGCAATATAAAAANGTGT	1080
Db	1021	GTTTCATGATAGGGGATGATTTTTCCTCCATTAAACAGAAATGCAATATAAAAANGTGT	1080
QY	1081	TTTTAAAAAAA 1094	
Db	1081	TTTTAAAAAAA 1094	
RESULT 2	AF151978	4520 bp	linear PRI 18-AUG-1999
LOCUS	AF151978	Homo sapiens amino acid transporter B0+ (ATB0+)	mRNA, complete cds.
DEFINITION	AF151978	GI:5732679	
ACCESSION	AF151978.1		
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
SOURCE			

OY	564	GAGCTGCTGGTTTGTAATTAACGCCCTACCCCTTTGATTTGGAATTTATATCGTCATTCG	623
Db	1757	GAGCTTCTGGTTTGTCATCACTCTCTATTCGTGTGTGCAATTCCTAGCTGGACATCG	1816
OY	624	TGCATTTTCATAGACCACTAATTTATGGCCCAATTCATCACTCGTACTGGGGAGTTGCTTAC	683
Db	1817	TGAATTTTCAACAGACCTGATATTCTCATATTCATATCCGACTGGGGAGTTCTCTTG	1876
OY	684	GCCTGGCTATGATTTGTTTCGCACTTATTTGGATTTCCAAATATGGCTATCAATAAAATTA	743
Db	1877	GCTGTGTATGATCATTTTCTGCATTTTGGATACCAATATATGGCTATCATATAAAATAG	1936
OY	744	TTACAGGCTAAAGAAACATCTTTCACAGCCCTTATAGTTGCTGCAGACCACTTCTTACT	803
Db	1937	TTACAGCGAGAAGAAACATACTCTTCAGGCTATCATATAGCTGCTGTAGACACACTCTCGA	1996
OY	804	GGGGTCCATACCTGGCAACACATCTGGGGGAAATATATTAAGACATGTAAGTCTCTAAA	863
Db	1997	GGGGTCCATACCTGGCAACACATCTGGGGGAAATATATTAAGACATGTAAGTCTCTAAA	2056
OY	864	AAGAGGCTGGCCATGAATAATACCTACTGTTTATGTGCAGACAGAAAACCGGAATGAGATCTCA	923
Db	2057	AGGAGACTGACCACGAATAATACCTACTGTTTATGTGCAGACAGAAAACCGGAATGAGATCTCA	2116
OY	924	---TTGAAAAAATATATGATTTGTAATATGATGATTTTTTTTATAGATATAGGGGACCCCTAT	980
Db	2117	CCCTTGGAAAAATCTAATTTGTAATATATATGATTTGTATAGATATAGGGGACATTTTAA	2176
OY	981	TTATTTTGTGTAACTGATAGATAGAAAATGACATACCTATGTCATGATAGAGT---GA	1036
Db	2177	TTATTTGTATGTATATCAATATAGGGAATATGACATGTCATCTCAGTCATGATGTATGA	2236
OY	1037	TTTTTTTCCCATTTAAGCAGAAATGCAATATATAAAATGTGGTTTTTTTA	1085
Db	2237	TACTTTTTCACCTTATATCAAGATAGCAATATATAAATGTGATCATATCA	2285

RESULT 4	AF161714	2337 bp	mRNA	linear	ROD 15-AUG-1999
LOCUS	AF161714				
DEFINITION	Mus musculus amino acid transporter B0+ (ATB0+)		mRNA, complete cds.		
ACCESSION	AF161714				
VERSION	AF161714.1	GI:5732879			
KEYWORDS					
SOURCE					
ORGANISM	Mus musculus (house mouse)				
	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 2337)				
AUTHORS	Sloane,J.L. and Mager,S.				
TITLE	Cloning and functional expression of a human Na(+) and				
	Cl(-)-dependent neutral and cationic amino acid transporter B(0+)				
JOURNAL	J. Biol. Chem. 274 (34), 23740-23745 (1999)				
REFERENCE	99376996				
PUBMED	10446133				
	2 (bases 1 to 2337)				
AUTHORS	Reweil,L., Sloane,J.L. and Mager,S.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-JUN-1999) Cell and Molecular Physiology, University				
	of North Carolina, CB 7545, 460 MSRB, Chapel Hill, NC 27599, USA				
FEATURES	Location/Qualifiers				
source	1..2337				

gene	1. .2337
	/gene="ATB0+"
CDS	249. .2165
	/gene="ATB0+"

```

/note="similar to the Homo sapiens ATB0+ encoded by the
sequence presented in GenBank Accession Number AF151978"
/codon_start=1
/product="amino acid transporter B0+"
/protein_id="A0A09320.1"
/db_xref="GI:5732880"
/translation="MDRLKCPNFFCKRQKERTVASSENFVHGENDEOERNKSKSD
YLSMWGAVAGLGNNWRRPYLLTYINGGAALIPALMALAGLPVLPFESLCOOFATL
GPVSWRIILPDLFOVGGITWVPIVSFVAIYVNIIVLIPALVAFSPOSLPWANCSMAA
DENOSRPIVATGCNVNVSIGAGEEMNINISVNTNNITCNGSEVERPQGLPEBOWDKYPT
LORSQDEGTGVVWVYALICLLAMLLIIVGALPEFSGSRRVYFTLIPFYVYLLIL
INRGKTEGASGKISIVYGAOSNFTKLEAEWKNAAQOIVSYSAACGAGVLAISYKSK
FNCCYEDAIIVLCITKCTISVFAGSFGATFISGHNAHISGKESVYSGSGDFAIAP
EALLOLPAGPEFASILPEFFMLITLGDPCQFASIIETITTFODLEFRAKRRNRVPTLDC
CLITFLILGLCVTQAGIIVYHMLIDBHCOPAGMIIAALITLGAITIMVIGNRFIDIEIM
MGCKRWKIPFIMLRACGCFVITPILLSAIIVMSIVTKPRPYADIIPYDWCVALQCMKAI
DIECIWTIPMILIRKIVQAGNIIILRIITISCCRPASNMWPEYLEKHNGERYRMAEPAKET
DHETPLTSGSKTPE"

```

[illegible]

QY 981 TTTATTTGTGTAACTAGTAAGAAATGATACATATGTCATGATAGGCT----GA 1036
 Db 2233 TTTATTTGTATGATTAATCAATTAGGAAATGTATCATGTCATGTCATGATGTTAG 2292
 QY 1037 TTTTATTTTCCATTTAGAGCAATCATATTAATAATGTGCTTTT 1081
 Db 2293 TACTTTTTCCTACTTAATCAAGAAATGAATATTAATTAATCTGAATCT 2337
 RESULT 5
 AB033285 1923 bp mRNA linear ROD 08-APR-2000
 LOCUS M1833285
 DEFINITION Mus musculus mCATB0+ mRNA for colonic system B0+ amino acid transporter, complete cds.
 ACCESSION AB033285
 VERSION AB033285.1 GI:7544119
 KEYWORDS Colonic system B0+ amino acid transporter.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1923)
 AUTHORS Ugaawa, S.
 JOURNAL Mouse Colonic System B0+ Amino Acid Transporter
 REFERENCE 2 (bases 1 to 1923)
 AUTHORS Ugaawa, S.
 JOURNAL Direct Submission
 Submitted (05-OCT-1999) Shinya Ugaawa, Nagoya City University Medical School, Department of Anatomy II, 1 Kawasumi, Mizuno-cho, Mizuno-ku, Nagoya, Aichi 467-0816, Japan
 (E-mail: ugaawa@med.nagoya-cu.ac.jp, Tel: 81-52-853-8126 (ex. 8126), Fax: 81-52-853-8887)
 FEATURES
 source
 1. 1923
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="ddy"
 /db_xref="taxon:10090"
 1. 1923
 /gene="mCATB0+"
 7. 1923
 /gene="mCATB0+"
 /codon_start=1
 /product="colonic system B0+ amino acid transporter"
 /protein_id="BA04300.1"
 /db_xref="GI:7544120"
 /translation="MDRLKCPNFCKROKKEVYASSPNHVGENDENQENSKSD YLSMVGAVGLGNVRFPLTYTNGGAFLIYATIMLALAGLPFLPESLQFNSL GPVSVRIILPFGVGTPTLVISVPAIYINATISLTLFASFOVLPMWNCSSWA DENCSPRTYTCNVSTIGAGEMFNISWNTNLTGNSSEVRPQLPSEQYMDVLT LKNSGMDGVVWVWYALCLLAMLIVGALFKGKSSGKVVYFALPVLVLLIL LKATLEGASRGKSYIYGAOSNFTKLREAVMDATQIFSLSVAMGGLVALSYNK FNNNCSDAIIVCLNCLSVFAGPAIFSLGMAHISGEVSOVYSGPDLAFIYV EALALQAPAGFNFSLFEFMLTLGLDSOFASITITTTTODLPKMKMRVPTTGC CLILFLGLCTVAGIYVWHLIDHCGAGGIIAALIEIAGITWITGGRITIEDIM MIKARKMILWLMWRACRPVITPPLLSAILVMSLVKTHRPDYADIPYDGVALGWCMI TFCITNIPIMATIKIVQABESNIIQRITISCRPASNMGPYLEKHGRBYRDMAPAKET DHEIPITSGSTKE"
 BASE COUNT 503 a 361 c 445 g 614 t
 ORIGIN
 Query Match 46.2%; Score 505.8; DB 10; Length 1923;
 Best Local Similarity 85.9%; Pred. No. 1, le-98;
 Matches 561; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
 QY 264 CCTCTGAGAAACGATACACAACAATTCAGATTATTTCCCAAGATGATGAAGAA 323
 Db 1271 CTCTCTATGAAACAATTTACACAACAATTCAGATTGTTTCCCAAGATGATGAAGAA 1330
 QY 324 TGAAGGTTCCCATTAACCTTGGCTGCTGCTTTGCTTTCTCTTGGCTCTGCTG 383
 Db 1331 TGAAGGTTCCCATTAACCTTGGCTGCTGCTTTGCTTTCTCTTGGCTCTGCTG 1390

QY 384 TGACTCAGCGTGAATTTACTGGGTGATCTGATTTGACACACTTCTGCTGATGGGCA 443
 Db 1391 TGACTCAGCGTGAATTTACTGGGTGATCTGATTTGATGATGATCTGCTGATGGGCA 1450
 QY 444 TTTTAAATTCAGCATATCTGAGCTACTTGGATCATCTGATTTATGAGGACAGAT 503
 Db 1451 TTTTAAATTCAGCATATCTGAGCATATCTGAGCATATCTGATTTATGAGGACAGAT 1510
 QY 504 TCATTGAGATACAGAAATGATGATTTGAGCAAGAGAGTGGATATCTGCTAATGGA 563
 Db 1511 TCATTGAGATATGAGAAATGATGATTTGAGCAAGAGAGTGGATATCTGCTAATGGA 1570
 QY 564 GAGCTTCTGCTGTTGTAATTAACGCTTATCTTTGATGCAATATTAATCTGCTATGG 623
 Db 1571 GAGCTTCTGCTGTTGTAATTAACGCTTATCTTTGATGCAATATTAATCTGCTATGG 1630
 QY 624 TGCATATTCATAGACCTAATTAATGAGCAATTCATACCTGAGCTGGAGGCTTTCAG 683
 Db 1631 TGCATATTCATAGACCTGATTAATGAGCAATTCATACCTGAGCTGGAGGCTTTCAG 1690
 QY 684 GCTGCTGATGATTTGTTCTGCAATATTTGATTCATATTCATATTCATATTCATATTA 743
 Db 1691 GCTGCTGATGATTTGTTCTGCAATATTTGATTCATATTCATATTCATATTCATATTA 1750
 QY 744 TTCAGGCTAAGAAACATCTTTCACGCTTATTAAGTCTGACAGACGCTTCTACT 803
 Db 1751 TTCAGGCTAAGAAACATCTTTCACGCTTATTAAGTCTGACAGACGCTTCTACT 1810
 QY 804 GGGGTCATACCTGTAACACATCTGAGGAAATATTAAGACATGATGATGCTCTAATA 863
 Db 1811 GGGGTCATACCTGTAACACATCTGAGGAAATATTAAGACATGATGATGCTCTAATA 1870
 QY 864 AAGAGCTGCGCCATGAATATCTGATGATGATGATGATGATGATGATGATGATGAT 916
 Db 1871 AAGAGCTGCGCCATGAATATCTGATGATGATGATGATGATGATGATGATGATGAT 1923
 RESULT 6
 HS452H17 99682 bp DNA linear PRI 05-MAR-2003
 LOCUS Human DNA sequence from clone RP3-452H17 on chromosome Xq22.1-23, complete sequence.
 ACCESSION 296810
 VERSION 296810.1 GI:2276315
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 99682)
 AUTHORS Grafham, D.
 JOURNAL Direct Submission
 Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Jul 24, 1997 this sequence version replaced gi:2198470.
 COMMENT
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep. This sequence was generated from part of bacterial clone configs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChX>. RP3-452H17 is from the library RPc1-3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

source

1.99682
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="q22.1-23"
/clone="RP3-452H17"
/clone_lib="RPc1-3"

BASE COUNT 28716 a 17592 c 17583 g 35791 t

ORIGIN

Query Match 35.8%; Score 392; DB 9; Length 99682;
Best Local Similarity 100.0%; Pred. No. 2e-74; 0; Indels 0; Gaps 0;
Matches 392; Conservative 0; Mismatches 0;

QY 1 GTTCACAGGGGAGCTGTACCTTACAGTTGTTATGATGAAAAATCATATTAACACACTA 60
DB 4240 GTTCACAGGGGAGCTGTACCTTACAGTTGTTATGATGAAAAATCATATTAACACACTA 4299
QY 61 AAAACAGCAATAGCGCTGATTTGATTTCAACAGAACCAAGGTGATGCGCAAAAGA 120
DB 4300 AAAACAGCAATAGCGCTGATTTGATTTCAACAGAACCAAGGTGATGCGCAAAAGA 4359
QY 121 AAAACCGCTGTGGGTTTTTAAAGCAAGAAAGGCGCTCAGACTATATTTTCAACATATA 180
DB 4360 AAAACCGCTGTGGGTTTTTAAAGCAAGAAAGGCGCTCAGACTATATTTTCAACATATA 4419
QY 181 TTGTGAATTTGATTAACATACATGTAATCTGTGATTAACAGTAATGGCAATTAAGTTG 240
DB 4420 TTGTGAATTTGATTAACATACATGTAATCTGTGATTAACAGTAATGGCAATTAAGTTG 4479
QY 241 ACATATTAACATGATCTTCTCCCTCTGAAGAAACGATCAACAACAATTCAGAATTT 300
DB 4480 ACATATTAACATGATCTTCTCCCTCTGAAGAAACGATCAACAACAATTCAGAATTT 4539
QY 301 ATTTCCCAAGATGATGAAGAAATGAGGGTTCCCATTAATTTGGGCTGCTGTTGTTT 360
DB 4540 ATTTCCCAAGATGATGAAGAAATGAGGGTTCCCATTAATTTGGGCTGCTGTTGTTT 4599
QY 361 GTTTCCTGCTGCTGCTGCTGCTGATGACAG 392
DB 4600 GTTTCCTGCTGCTGCTGCTGCTGATGACAG 4631

RESULT 7

BX119904/c

LOCUS BX119904 178704 bp DNA linear PRI 10-FEB-2003
DEFINITION Human DNA sequence from clone RP11-232D9 on chromosome X, complete
Sequence.

ACCESSION BX119904
VERSION BX119904.4 GI:28193290

KEYWORDS HTG.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 178704)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Howden, P.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Feb 1, 2003 this sequence version replaced gi:27902176.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep. This sequence was generated from part of bacterial clone configs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChX>.

RP11-232D9 is from the library RPc1-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

FEATURES

source

1.178704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-232D9"
/clone_lib="RPc1-11.1"

BASE COUNT 56317 a 32304 c 35229 g 56554 t

ORIGIN

Query Match 30.6%; Score 334.4; DB 9; Length 178704;
Best Local Similarity 90.8%; Pred. No. 4.6e-62;
Matches 356; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 GTTCACAGGGGAGCTGTACCTTACAGTTGTTATGATGAAAAATCATATTAACACACTA 60
DB 65147 GTTCACAGGGGAGCTGTACCTTACAGTTGTTATGATGAAAAATCATATTAACACACTA 65088
QY 61 AAAACAGCAATAGCGCTGATTTGATTTCAACAGAACCAAGGTGATGCGCAAAAGA 120
DB 65087 AAAACAGCAATAGCGCTGATTTGATTTCAACAGAACCAAGGTGATGCGCAAAAGA 65028
QY 121 AAAACCGCTGTGGGTTTTTAAAGCAAGAAAGGCGCTCAGACTATATTTTCAACATATA 180
DB 65027 AAAACCGCTGTGGGTTTTTAAAGCAAGAAAGGCGCTCAGACTATATTTTCAACATATA 64968
QY 181 TTGTGAATTTGATTAACATACATGTAATCTGTGATTAACAGTAATGGCAATTAAGTTG 240
DB 64967 TTGTGAATTTGATTAACATACATGTAATCTGTGATTAACAGTAATGGCAATTAAGTTG 64908
QY 241 ACATATTAACATGATCTTCTCCCTCTGAAGAAACGATCAACAACAATTCAGAATTT 300

```

Db      64907 AATATATACATGCTATTTTCCCTCGAAGAGATCAAAACAGATTCAAGATT 64848
QY      301 ATTTCCCAAGATGAGAAATAGAGTTCCCACTTGGCTGCTGCTGCTTT 360
Db      64847 ACTTCCCAAGATGAGAAATAGATGATGCTTCAACTTGGCTGCTGCTTT 64788
QY      361 GTTCTCCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392
Db      64787 GTTCTCCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 64756

RESULT 8
AC133215
LOCUS   AC133215 151973 bp DNA linear HTG 07-SEP-2002
DEFINITION Homo sapiens chromosome UNK clone RP13-792C11, WORKING DRAFT
SEQUENCE AC133215
VERSION  AC133215.1 GI:22758609
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          Waterston,R.H.
          The sequence of Homo sapiens clone
          Unpublished
          2 (bases 1 to 151973)
          Waterston,R.H.
          Direct Submission
          Submitted (07-SEP-2002) Genome Sequencing Center, Washington
          University School of Medicine, 4444 Forest Park Parkway, St. Louis,
          MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
          Project Information
          Center project name: H.FH0792C11
----- Summary Statistics -----
Sequencing vector: M13; 08
Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 149763 bases at least Q40
Consensus quality: 149763 bases at least Q30
Consensus quality: 149919 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1 1119: contig of 1119 bp in length
* 1120 1219: gap of unknown length
* 1220 2317: contig of 1098 bp in length
* 2318 2417: gap of unknown length
* 2418 4383: contig of 1966 bp in length
* 4384 4484: gap of unknown length
* 4484 6519: contig of 2036 bp in length
* 6520 10515: gap of unknown length
* 10516 10615: contig of 3896 bp in length
* 10616 18127: gap of unknown length
* 18128 18227: gap of unknown length

```

```

* 18228 29110: contig of 10883 bp in length
* 29111 29210: gap of unknown length
* 29211 47743: contig of 18533 bp in length
* 47744 47844: gap of unknown length
* 47844 69544: contig of 21701 bp in length
* 69545 92308: gap of unknown length
* 92309 92408: contig of 22664 bp in length
* 92409 123518: gap of unknown length
* 123519 123619: contig of 3110 bp in length
* 123619 151973: contig of 28355 bp in length.

FEATURES
source
1. 151973
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="UNK"
   /clone="RP13-792C11"
1. 1119
   /note="assembly_name:Contig1"
1220. 2317
   /note="assembly_name:Contig5"
2418. 4383
   /note="assembly_name:Contig8"
4484. 6519
   /note="assembly_name:Contig9"
6620. 10515
   /note="assembly_name:Contig10"
10616. 18127
   /note="assembly_name:Contig11"
18228. 29110
   /note="assembly_name:Contig12"
29211. 47743
   /note="assembly_name:Contig13
clone_end:17
vector_side:left"
47844. 69544
   /note="assembly_name:Contig14"
69645. 92308
   /note="assembly_name:Contig15"
92409. 123518
   /note="assembly_name:Contig16"
123619. 151973
   /note="assembly_name:Contig17
clone_end:SP6
vector_side:left"

BASE COUNT  47782 a 28922 c 29161 g 45007 t 1101 others
ORIGIN
Query Match      28.0%; Score 305.8; DB 2; Length 151973;
Best Local Similarity 76.1%; Pred. No. 6,8e-56;
Matches 451; Conservative 0; Mismatches 117; Indels 25; Gaps 5;

QY      375 TCGTGTGTGATCAGGCTGGAATTTACTGGGTCATGCTGATGACCACTTGTGCTG 434
Db      146271 TCACCGGTGTGCTGAGCTGGAATTTACT -GGTCATGCTGACCTGGCACTTGTGCTG 146329
QY      435 GATGGGCGATTTAATGACGCTATACCTGAGCTAGTTGAATCATCTGGATTATGAG 494
Db      146330 AACGGACATTTGATGCAACTATACCTGGAATTAAGAAATCATCTGGATTATGAG 146389
QY      495 GGACAGATTCATTTGATGACAGAAATGATGATGGAGCAAGAGGTGATATTTCTGGC 554
Db      146390 GGACAGATTCCTTGAAGATATAAATATCATATAT ---GGACAGAGGTGATATTTCTGGC 146446
QY      555 TATGTGGAGACCTTGCTGCTTGTG-----AATTAGCCTATCCTTTTG 598
Db      146447 TCACAGTGAAGAACTGCTGCTTGTGCTATTAATGATGACATTAATGCTATATTTT 146506
QY      599 ATTGCAATTTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
Db      146507 ATTGCAATTTTATCTGCACTAGTGAATATATCATTAAGCTTGAATTATGCAAAATCTTG 146566

```

```

OY      659  TACCCGACTGGGAGTGGCTTACGCTGGTGTATGATGTTTCTGCATTATTTGGATT 718
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      146567 TCCCCTGACTGAGAGTGTCTCAAGGCTGTGTATGATTAATTTCTGTATTA-TTGCATT 146625
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      719  CCAATTATGCTATATATAAATAATATTCAGGCTTAAGGAAACATCTTCAACGCTTTATA 778
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      146626 TCAATTATGCTATATATAAATAATATTCAGGCTTAAGGAAACATCTTCTGTGTGTATA 146685
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      779  AGTGTCTGCAGACGACCTTCTTAACCTGGGCTCCATACCTGGAACAACATCTGGGAGAGA 838
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      146686 AAGCTGTGTATTAAGAGTTTTCATAGACCCATACCTGGAAGACATCATGCGGAGAAA 146745
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      839  TATTAAGACATGTGACTTCTTAATAAAGAGG-----CTGGCCATGAAATACCTACTGTAG 894
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      146746 TAAAAAGGACAGATGATCTTAAATTAATAAAGAGACTTGCAGGAATTCCTACTGTGG 146805
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      895  TGGCAGCAGAAACCGAATGATCATCTGTAATAAATAATATATGTTGTATA 947
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      146806 TGGCAGAGGAAACCGAATGATCATCTGTAATAAATAATGTTGTATA 146858
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 9
AC023347 174655 bp DNA linear PRI 09-MAY-2001
LOCUS Homo sapiens BAC clone RP11-507H23 from 2, complete sequence.
DEFINITION AC023347
AC023347.8 GI:11597167
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

REFERENCE 1 (bases 1 to 174655)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 174655)
AUTHORS Swearengen, S., Maupin, R. and Barrett, M.
TITLE The sequence of Homo sapiens BAC clone RP11-507H23
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 174655)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 174655)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 174655)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 6 (bases 1 to 174655)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Dec 7, 2000 this sequence version replaced gi:9690405.

```

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 Summary Statistics
 Center project name: H_NH0507H23

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RP11-507H23 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatem, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6
 NEIGHBORING SEQUENCE INFORMATION:
 Actual start of this clone is at base position 1 of RP11-507H23; actual end is at base position 174655 of RP11-507H23.

```

FEATURES
Source
1..174655
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-507H23"
/clone_1kb="RP11-11"
1329..2040
/repeat_region
/rpt_family="L1"
2040..2073
/repeat_region
/rpt_family="L1"
3468..3639
/repeat_region
/rpt_family="MIR"
5493..5717
/rpt_family="Alu"
10078..10146
/rpt_family="MER94"
10148..10617
/repeat_region
/rpt_family="L2"
11561..11711
/rpt_family="MIR"
13756..13814
/rpt_family="CR1"
13816..14212
/repeat_region
/rpt_family="ERV1"
14359..14613
/rpt_family="L1"
14614..14908
/repeat_region
/rpt_family="Alu"
14909..15067
/repeat_region
/rpt_family="L1"
15078..15138
/repeat_region
/rpt_family="L1"
15816..16112
/repeat_region
/rpt_family="L1"
16202..16252
/repeat_region
/rpt_family="L2"

```

```
repeat_region 17454. .18620 /rpt_family="L1"
repeat_region 18777. .19121 /rpt_family="MALR"
repeat_region 19259. .19367 /rpt_family="MALR"
repeat_region 20344. .20420 /rpt_family="CRL"
repeat_region 20595. .20900 /rpt_family="MIR"
repeat_region 21011. .21130 /rpt_family="Alu"
repeat_region 22367. .22439 /rpt_family="MIR"
repeat_region 22440. .22728 /rpt_family="MALR"
repeat_region 22729. .23277 /rpt_family="Alu"
repeat_region 23300. .23387 /rpt_family="MALR"
repeat_region 23388. .23785 /rpt_family="MALR"
repeat_region 23786. .23882 /rpt_family="L2"
repeat_region 24630. .24733 /rpt_family="L1"
repeat_region 24970. .25027 /rpt_family="L1"
repeat_region 25974. .26806 /rpt_family="L2"
repeat_region 26832. .27067 /rpt_family="L1"
repeat_region 28690. .29052 /rpt_family="L1"
repeat_region 29118. .29663 /rpt_family="L2"
repeat_region 30891. .31010 /rpt_family="L1"
repeat_region 31011. .31308 /rpt_family="L1"
repeat_region 31309. .31350 /rpt_family="Alu"
repeat_region 31351. .31716 /rpt_family="L1"
repeat_region 31717. .31830 /rpt_family="MERL_type"
repeat_region 32007. .32287 /rpt_family="L1"
repeat_region 32341. .32413 /rpt_family="L1"
repeat_region 32453. .32998 /rpt_family="L1"
repeat_region 33534. .33712 /rpt_family="ERV1"
repeat_region 33771. .34069 /rpt_family="MERS3"
repeat_region 34517. .34604 /rpt_family="Alu"
repeat_region 34848. .34980 /rpt_family="L1"
repeat_region 34981. .35528 /rpt_family="L1"
repeat_region 35628. .35872 /rpt_family="L1"
repeat_region 35873. .36922 /rpt_family="L1"
repeat_region 36923. .37302 /rpt_family="L1"
repeat_region 37678. .37925 /rpt_family="L1"
repeat_region 37939. .38108 /rpt_family="MIR"
repeat_region 39114. .39713 /rpt_family="MERL_type"
```

```
repeat_region /rpt_family="L1"
repeat_region 39707. .39988 /rpt_family="L1"
repeat_region 40613. .40725 /rpt_family="L1"

Query Match 28.0%; Score 305.8; DB 9; Length 174655;
Best Local Similarity 76.1%; Pred. No. 6.6e-56;
Matches 451; Conservative 0; Mismatches 117; Indels 25; Gaps 5;

QY 375 TCGTCGTGACACGAGCTGGAAATTTACTGAGTTCATCTGATGACCACTTCGTGCTG 434
DB 122834 TCACCGCTGTGTCAGAGCTGGAAATTTAGT-GGCTCATCTGACGCGCACTTCGTGCTG 122776
QY 435 GATGGGCAATTTTATTTATTCAGCTTACTGAGAGTGTGAGATCATGATTTATGAG 494
DB 122775 AACGGACATTTTATTTATTCAGCTTACTGAGATCATTTAGATCATCTGGTTTATGAG 122716
QY 495 GGAACAGATTCATTGAGATACAGAAATGATGATTTGAGCAAAAGGTGATTTATTCGCG 554
DB 122715 GGAACAGATTCATTGAGATACAGAAATGATGATTTGAGCAAAAGGTGATTTATTCGCG 122659
QY 555 TATGCTGAGAGCTTCGTGCTTGT-----AATTACGCTTATCCTTTTG 598
DB 122658 TGCAGTGAACACTTGTGCTGTTGTCATTAATGCTATATGACATTTATGCTTATTTT 122599
QY 599 ATTGCATATTTTATTCGTCATGTCGCAATTTGATGACCTATTTATGGGCAATTCGA 658
DB 122598 ATTGCATATTTTATTCGTCATGTCGCAATTTGATGACCTATTTATGGGCAATTCGA 122539
QY 659 TACCTGACCTGGGAGATTTGCTTTAGGCTGTGATGATTTGCTTATTTGATTTGATTT 718
DB 122538 TGCCTGACCTGAGAAATTTGCTTCAAGGCTGTGATGATTTGCTTATTTGATTTGATTT 122480
QY 719 CCAATTATGCTATTCATTAATAATATTCAGCTTAAGAAACATCTTCAAGCCTTATA 778
DB 122479 TCAATTATGCTATTCATTAATAATATTCAGCTTAAGAAACATCTTCAAGCCTTATA 122420
QY 779 AGTTGCTGACAGACGCTTCTTAACGAGTCCATTCATTCGAGACAACTGCGGAGAAAGA 838
DB 122419 AGCTGCTGTAACAGATTTCAACCTAGAGCCATTCCTGAGAGACATCATGGGAGAAA 122360
QY 839 TATTAAGACATGTAGTCTTAAATAAGAGC---CTGCCATGAATTAACCTACTGTAG 894
DB 122359 TAAATAGGACAGTATGATCTTAAATTAAGAGCCTTGAACAGAAATTCCTACTGTGG 122300
QY 895 TGGCAGCAGAAACCGGAATGAGATTCATGTAATAATAATGATTTGATTA 947
DB 122299 TGGCAGCAGAAACCGGAATGAGATTCATGTAATAATAATGATTTGATTA 122247

RESULT 10
AC027640 158527 bp DNA linear HTG 01-SEP-2000
LOCUS Homo sapiens chromosome X clone RP11-714C1, WORKING DRAFT SEQUENCE,
DEFINITION 33 unordered pieces.
ACCESSION AC027640
VERSION AC027640.3 GI:9958776
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 158527)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL unpublished
REFERENCE 2 (bases 1 to 158527)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Sep 1, 2000 this sequence version replaced gi:8318566.
```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0714C01
Summary Statistics
Sequencing vector: M13, 100%
Sequencing vector: plasmid, 0%
Chemistry: Dye-primer ET, 100% of reads
Chemistry: Dye-terminator Big Dye, 0% of reads
Assembly program: Phrap, version 0.990319
Consensus quality: 148926 bases at least Q40
Consensus quality: 151403 bases at least Q40
Consensus quality: 152420 bases at least Q20
Insert size: 155327; sum-of-contigs
Quality coverage: 4.92 in Q20 bases; sum-of-contigs
Quality coverage: 5.23 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
3335: contig of 3335 bp in length
3336: gap of unknown length
3436: contig of 2401 bp in length
5837: gap of unknown length
5837: gap of unknown length
8911: contig of 2975 bp in length
9012: contig of 2741 bp in length
11752: contig of 2741 bp in length
11852: gap of unknown length
14013: contig of 2161 bp in length
14014: gap of unknown length
14114: contig of 3165 bp in length
17278: gap of unknown length
17378: gap of unknown length
17379: contig of 2112 bp in length
19490: gap of unknown length
19591: contig of 2238 bp in length
21828: gap of unknown length
21928: gap of unknown length
24934: contig of 3006 bp in length
25034: gap of unknown length
25035: contig of 3285 bp in length
28319: gap of unknown length
28420: contig of 3163 bp in length
31582: gap of unknown length
31683: contig of 2976 bp in length
34658: gap of unknown length
34759: gap of unknown length
38324: contig of 3566 bp in length
38325: gap of unknown length
38424: gap of unknown length
41243: contig of 2819 bp in length
41343: gap of unknown length
41344: gap of 2203 bp in length
43546: gap of unknown length
43646: gap of unknown length
46071: contig of 2425 bp in length
46171: gap of unknown length
63170: contig of 16999 bp in length
63270: gap of unknown length
63271: gap of 4293 bp in length
67563: gap of unknown length
67564: gap of unknown length
72467: gap of 4804 bp in length
72568: gap of unknown length
77552: contig of 4985 bp in length
77553: gap of unknown length
83764: contig of 6112 bp in length
83864: gap of unknown length
90080: contig of 6216 bp in length
90180: gap of unknown length

```

```

FEATURES
source
1..158527
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-714C1"
1..3335
/feature="assembly_name:Contig10"
/feature="assembly_name:Contig10"
3436..5836
/feature="assembly_name:Contig11"
/feature="assembly_name:Contig11"
5937..8911
/feature="assembly_name:Contig12"
/feature="assembly_name:Contig12"
9012..11752
/feature="assembly_name:Contig13"
/feature="assembly_name:Contig13"
11853..14013
/feature="assembly_name:Contig14"
/feature="assembly_name:Contig14"
14114..17278
/feature="assembly_name:Contig15"
/feature="assembly_name:Contig15"
17379..19490
/feature="assembly_name:Contig16"
/feature="assembly_name:Contig16"
19591..21828
/feature="assembly_name:Contig17"
/feature="assembly_name:Contig17"
21929..24934
/feature="assembly_name:Contig18"
/feature="assembly_name:Contig18"
25035..28319
/feature="assembly_name:Contig19"
/feature="assembly_name:Contig19"
28420..31582
/feature="assembly_name:Contig20"
/feature="assembly_name:Contig20"
31683..34658
/feature="assembly_name:Contig21"
/feature="assembly_name:Contig21"
34759..38324
/feature="assembly_name:Contig22"
/feature="assembly_name:Contig22"
38425..41243
/feature="assembly_name:Contig23"
/feature="assembly_name:Contig23"
41344..43546
/feature="assembly_name:Contig24"
/feature="assembly_name:Contig24"
43647..46071
/feature="assembly_name:Contig25"
/feature="assembly_name:Contig25"
46172..63170
/feature="assembly_name:Contig26"
/feature="assembly_name:Contig26"
63271..67563
/feature="assembly_name:Contig27"
/feature="assembly_name:Contig27"
67564..72467
/feature="assembly_name:Contig28"
/feature="assembly_name:Contig28"
72568..77552
/feature="assembly_name:Contig29"
/feature="assembly_name:Contig29"
77553..83764
/feature="assembly_name:Contig30"
/feature="assembly_name:Contig30"
83864..90080
/feature="assembly_name:Contig31"
/feature="assembly_name:Contig31"
90081..90180
/feature="assembly_name:Contig32"
/feature="assembly_name:Contig32"

```

```

misc_feature      /note="assembly_name:Contig29"
83865..90080
/note="assembly_name:Contig30"
misc_feature      90181..97167
/note="assembly_name:Contig31"
misc_feature      97268..106868
/note="assembly_name:Contig32"
misc_feature      106969..114304
/note="assembly_name:Contig33"
misc_feature      114405..129608
/note="assembly_name:Contig34"
misc_feature      129709..146858
/note="assembly_name:Contig35"
misc_feature      146959..148941
/note="assembly_name:Contig4"
misc_feature      149042..150952
/note="assembly_name:Contig5"
misc_feature      151053..152328
/note="assembly_name:Contig6"
misc_feature      152629..154322
/note="assembly_name:Contig7"
misc_feature      154423..156637
/note="assembly_name:Contig8"
misc_feature      156738..158527
/note="assembly_name:Contig9"

BASE COUNT      48002 a 28204 c 31062 g 48053 t 3206 others
ORIGIN
Query Match      22.5%; Score 246; DB 2; Length 158527;
Best Local Similarity 87.0%; Pred. No. 5,1e-43;
Matches 282; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY      765 TTCAACGCTTTAAAGTTCGTCGACAGCAGCTTCTAAGTGGGCTGCATACCTGGAACAC 824
DB      107845 TTCAAGTGCCTTGAAGTTCGTCGACAGCAGCTTCTAAGTGGGCTGCATACCTGGAACAC 107786

QY      825 ATCTGGGGAAAGATATTAAGCATGAGTTCCTCAAAAAGAGCGCTGGCCATGAATAC 884
DB      107785 GTTCGGGGAGAGATATTAAGCATGATATATCCCAAAAAGAGCATGACATACATAC 107726

QY      885 CTACTGTTAGTGGCAGCAAGAAACCGAATGAGATCTCATTTGAAAAAAATATATGATGTT 944
DB      107725 CTACTGTTAGTGGCAGCAAGAAACCGAATGAGATCTCATTTGAAAAAAATATATGATGTT 107666

QY      945 ATAATGATGAT--TTTTTAAATAGGAGGACCTTATTTATTTGTTGTTAACTGATA 1002
DB      107665 ATTAATGTTATGTTTATTAAGAAATGGGGAACCTTTTATTTGTTATGTTAACTGATA 107606

QY      1003 GGAATATGATACATATGATGATGATGATGATTTTCCCATTTAGCAGGATGC 1062
DB      107605 GGAATATGATACATATGATGATGATGATTTTCCCATTTAGCAGGATGC 1062

QY      1063 AATATAAAAATGCTGTTTTTTAA 1086
DB      107545 AATATAAAAATGCAATTTCTTAA 107522

RESULT 11
LOCUS      BX119904      178704 bp      DNA      linear      PRI 10-FEB-2003
DEFINITION Human DNA sequence from clone RP11-232D9 on chromosome X, complete
ACCESSION   BX119904
VERSION     BX119904.4 GI:28193290
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 178704)
AUTHORS    Howden,P.
TITLE      Direct Submission
JOURNAL    Submitted (06-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,

```

```

COMMENT
Cambridgehire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 1, 2003 this sequence version replaced gi:27902176.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
EM, EMBL; SW, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/c.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chx
RP11-232D9 is from the library RPc1-11.1 constructed by the group
of Pletier de Jong. For further details see
http://www.chori.org/Dacpac/home.htm
VECTOR: pBAC3.6.

FEATURES
source
location/Qualifiers
1..178704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-232D9"
/clone_1lb="RPc1-11.1"

BASE COUNT      56317 a 32304 c 33529 g 36554 t
ORIGIN
Query Match      22.5%; Score 246; DB 9; Length 178704;
Best Local Similarity 87.0%; Pred. No. 5e-43;
Matches 282; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY      765 TTCAACGCTTTAAAGTTCGTCGACAGCAGCTTCTAAGTGGGCTGCATACCTGGAACAC 824
DB      83945 TTCAAGTGCCTTGAAGTTCGTCGACAGCAGCTTCTAAGTGGGCTGCATACCTGGAACAC 84004

QY      825 ATCTGGGGAAAGATATTAAGCATGAGTTCCTCAAAAAGAGCGCTGGCCATGAATAC 884
DB      84005 GTTCGGGGAGAGATATTAAGCATGATATATCCCAAAAAGAGCATGACATACATAC 84064

QY      885 CTACTGTTAGTGGCAGCAAGAAACCGAATGAGATCTCATTTGAAAAAAATATATGATGTT 944
DB      84065 CTACTGTTAGTGGCAGCAAGAAACCGAATGAGATCTCATTTGAAAAAAATATATGATGTT 94124

QY      945 ATAATGATGAT--TTTTTAAATAGGAGGACCTTATTTATTTGTTGTTAACTGATA 1002
DB      84125 ATTAATGTTATGTTTATTAAGAAATGGGGAACCTTTTATTTGTTATGTTAACTGATA 84184

QY      1003 GGAATATGATACATATGATGATGATGATTTTCCCATTTAGCAGGATGC 1062
DB      84185 GGAATATGATACATATGATGATGATGATTTTCCCATTTAGCAGGATGC 84244

QY      1063 AATATAAAAATGCTGTTTTTTAA 1086
DB      107545 AATATAAAAATGCAATTTCTTAA 107522

```



```

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc.help@bcm.tmc.edu
-----
Project Information
Center project name: KBM
Center clone name: CH230-460J22
-----
Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 186506 bases at least Q40
Consensus quality: 188708 bases at least Q30
Consensus quality: 190035 bases at least Q20
Estimated insert size: 191507; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 214592: contig of 214592 bp in length
* * 214593 214692: gap of unknown length
* * 214693 216341: contig of 1649 bp in length.
Location/Qualifiers
    1..216341
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-460J22"
            1..1124
                /note="ygs_contig"
                    misc_feature
                        207607..208981
                            /note="ygs_contig"
                                misc_feature
                                    213081..214592
                                        /note="ygs_contig"
                                            BASE COUNT      60786 a 35707 c 36312 g 58458 t 25078 others
                                                ORIGIN
Query Match          15.6%; Score 171.2; DB 2; Length 216341;
Best Local Similarity 74.2%; Pred. No. 6,2e-27;
Matches 245; Conservative 0; Mismatches 78; Indels 7; Gaps 2;

OY 763 CTTTCAACGCTTAATAAGTTGGTCGAGACCAGCTTCTAATCGGGGTCCATACCTGGAACA 822
    | ||||| |
Db 112155 CCTCAGCGCTATCATCAVAAAATGCTGTAGAACCAAGCTTCAAACACTGGGCTCATACCTGAAAA 1120966

OY 823 ACATCGTGCGGGAAGAAGTATTAAGACATGCTGATGTTCTTAATAAAANGCGTGCCTATGAAAT 882
    | ||||| |
Db 112095 ACATCTCGTAGAGAGAGATATTAAGACATGCTGAACCTGCCAAAGAGACTGACCCAGCAAT 1120366

OY 883 ACCCTACTGTTAAGGAGAGAGAAAACCGGAATGAGATCTCA---TTGAAAAAATAATATATG 939
    | ||||| |
Db 112035 ACCCTACTATTAGTGGCACCTAGAAAACCGAATAGACATCTCACTTTGAAAAATATCGAATT 1119766

OY 940 ATTGTATAATGATGATTTTTTTTGAATAGAGGGGACCTATTATTTGTGTGTTAACTGA 999
    | | ||||| |
Db 111975 GTTCTATATAGATGATTTGTGTAGAGTAGGGGGAATTTTAATTATTTGTATGTTAAACAA 1119166

OY 1000 ATNAGAAAATGTACATCACTATGTTATATAGGGT----GATTTTTTCCCATTTAAGCA 1055
    | |||| |
Db 111915 CTAGAGGGGAGTACTGTCCTCATTTAGTACAGGATGTAACTACTTTTACCACTTAATCA 1118566

OY 1056 GGATGCAATATATAAAATGCTGGTTTTTTTA 1085
    | |||| |
Db 111855 AGAATCAAAATGTTAAAAATGTGAATTTATCA 111826.

```

AC108660
 LOCUS 323793 bp DNA linear HMG 22-SEP-2002
 DEFINITION Rattus norvegicus clone CH230-74G23, *** SEQUENCING IN PROGRESS
 AC108660
 AC108660.4 GI:23101185
 HG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus

REFERENCE AUTHORS

Muzny, D., Marie, M., Lee, A., Adams, C., Alder, J., Allen, C., Allen, H., Albrechts, S., Amin, A., Angiano, D., Anyalebech, V., Ayag, A., Ayodeji, M., Bacc, E., Baden, H., Baldwin, D., Bandaru, D., Barber, M., Barnstead, J., Benahmed, F., Bivado, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Burch, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Y., Chen, Y., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Delgado, O., Denson, S., Deramo, C., Ding, Y., Din, H., Diya, R., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frisler, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., George, E., Geer, K., Gill, R., Gladys, A., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoque, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idler, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kover, C., Kowals, C., Kraft, C.L., Lebow, H., Levan, Z., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, D., Lorenshewe, L., Louised, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Marqu, P., Martin, K., Martin, R., Martinez, E., Marzley, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Miliomavljic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nait, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okunolu, G., Olajun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Plazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Sma, J., Speed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczek, R., Wooden, H., Worley, F., Wu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von Weizsacker, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

JOURNAL

COMMENT

Submitted (22-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Sep 18, 2002 this sequence version replaced g1:21737904.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence contigs will be indicated in the feature table.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GRYA

Center clone name: CH230-74G23

Summary Statistics

Assembly program: Phrap; version 0.990329
 Consensus quality: 202774 bases at least Q40
 Consensus quality: 205089 bases at least Q40
 Consensus quality: 20547 bases at least Q20
 Estimated insert size: 226383; sum-of-contigs estimation
 Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished genome as soon as it is available and the accession number will be preserved.

1 312024: contig of 312024 bp in length
 * 312025 312124: gap of unknown length
 * 312125 316461: contig of 4337 bp in length
 * 316462 316561: gap of unknown length
 * 316562 317688: contig of 1127 bp in length
 * 317689 317788: gap of unknown length
 * 317789 319338: contig of 1550 bp in length
 * 319339 319238: gap of unknown length
 * 319239 321271: contig of 2033 bp in length
 * 321272 321371: gap of unknown length
 * 321372 323793: contig of 2422 bp in length.

FEATURES

source

1. 323793

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-74G23"

/complement(1136..1658)

/note="clone_boundary"

clone_end:77

site:ECORI

end_sequence: BH37444"

192400..194119

/note="wgs_contig"

/complement(219624..220511)

/note="clone_boundary"

clone_end:Sp6

site:ECORI

end_sequence: BH37447"

BASE COUNT 65314 a 38957 c 38191 g 64287 t 116004 others

ORIGIN

Query Match 15.6%; Score 171.2; DB 2; length 323793;
 Best Local Similarity 74.2%; Pred. No. 5,9e-27;
 Matches 245; Conservative 0; Mismatches 78; Indels 7; Gaps 2;

```
OY 763 CTTTCAAGCCTTATAGTGTGCGACGAGCTTCTAATGCGGTCATACCTGGAACA 822
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74209 CTTTCAAGCCTTATAGTGTGCGACGAGCTTCTAATGCGGTCATACCTGGAACA 74268
OY 823 ACATCGTGGGGAAGATATAAAGACATGCTTCTTAAAAAGAGGCTGGCCATGAAAT 882
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74269 ACATCGTGGGGAAGATATAAAGACATGCTTCTTAAAAAGAGGCTGGCCATGAAAT 74328
OY 883 ACCTACTGTACTGGCAGAGAAAACGGAAATGAGATCTCA---TTGAAAAAAATATATG 939
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74329 ACCTACTGTACTGGCAGAGAAAACGGAAATGAGATCTCACTTTGAAAAATGCAAT 74388
OY 940 ATTGTATATGTGATTTTATAGAAATAGGGGACCTTATTTATTTGTGTACTGA 999
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74389 GTTCTATATGTGATTTTATAGAAATAGGGGACCTTATTTATTTATTTGATGTAATCAA 74448
OY 1000 ATAGGAAATGTACATACATATGTCATGATAGGGT---GATTTTTCCTTACGA 1055
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74449 CTAGGGGATGTACCTGCTCAGTACAGGATTTGTAAGTACTTTACCACTTAATCA 74508
OY 1056 GGAATGCATATPAAAAATGTGTTTATTA 1085
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74509 AGAATGAATGTAAAAATGGAATTAATCA 74538
```

Search completed: August 27, 2003, 11:20:46
Job time : 4380 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 17:23:25 ; Search time 85 Seconds

(Without alignments)
655,757 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198

Sequence: 1 SEETITTTIDLPFKVKKM.....VPKKAGHEIPVSGSRKPE 216

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_proteobacteria:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvivirus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1176	98.2	642	4	Q9UN76
2	1106.5	92.4	331	1	Q15003
3	1074	89.6	638	11	Q9JMA9
4	1074	89.6	638	11	Q91Y60
5	1074	89.6	638	11	Q9R183
6	1066	89.0	638	11	Q9D317
7	508.5	42.4	791	11	Q91XQ2
8	508.5	42.4	799	11	Q8CFM5
9	344.5	28.8	633	11	Q8VC47
10	332	27.7	631	5	Q9V7R0
11	332	27.7	631	5	Q9NB97
12	327	27.3	619	11	Q9U41
13	327	27.3	619	11	Q8C009
14	327	27.3	629	13	Q902V1
15	322	26.9	619	11	Q9R0X6
16	319	26.6	620	6	Q9GJ75

17	319	26.6	620	6	Q9GJ76	Q9GJ76 macaca fasc
18	317	26.5	620	6	Q9GJ77	Q9GJ77 salmtr1 scl
19	317	26.5	1201	5	Q9WJ10	Q9WJ10 drosophila
20	315	26.3	658	5	Q94917	Q94917 drosophila
21	315	26.3	943	5	Q9VR07	Q9VR07 drosophila
22	315	26.3	943	5	Q9VR07	Q9VR07 drosophila
23	311	26.0	674	5	Q8WPM9	Q8WPM9 olkopleura
24	307.5	25.7	599	4	Q8NKA8	Q8NKA8 homo sapien
25	305	25.5	635	13	Q9DCN5	Q9DCN5 gallus gall
26	303	25.3	630	13	Q42482	Q42482 rana catesb
27	302.5	25.3	629	13	Q9DE16	Q9DE16 oreochromis
28	300.5	25.1	622	13	Q91503	Q91503 torpeda mar
29	299.5	25.0	628	4	Q96KH8	Q96KH8 homo sapien
30	299	25.0	328	4	Q9BM84	Q9BM84 homo sapien
31	299	25.0	572	6	Q9MY7	Q9MY7 macaca mulle
32	298	25.0	617	6	Q9MY8	Q9MY8 macaca mulle
33	298	24.9	639	5	Q9VJ4	Q9VJ4 drosophila
34	298	24.9	739	5	Q9NKA8	Q9NKA8 drosophila
35	294.5	24.6	625	13	Q91BH6	Q91BH6 cyprinus ca
36	293.5	24.5	135	6	Q28498	Q28498 macaca mulle
37	293	24.5	617	11	Q8R212	Q8R212 mus musculu
38	292	24.4	567	11	Q9WTR3	Q9WTR3 rattus norv
39	292	24.4	597	11	Q9WTR4	Q9WTR4 rattus norv
40	292	24.4	598	13	Q73771	Q73771 raja sp. ga
41	292	24.4	617	11	Q63380	Q63380 rattus norv
42	290.5	24.2	625	13	Q8JHX8	Q8JHX8 salmo salar
43	290	24.2	374	5	Q962P1	Q962P1 aplysia cal
44	289	24.1	640	5	Q81S70	Q81S70 aedes aegypt
45	288	24.0	598	13	Q91494	Q91494 torpeda cal

ALIGNMENTS

RESULT 1

Q9UN76

PREDIMINARY;

PRT; 642 AA.

AC Q9UN76;

01-MAY-2000 (TREMUR1.13, Created)

01-MAY-2000 (TREMUR1.13, Last sequence update)

DT 01-OCT-2002 (TREMUR1.22, Last annotation update)

Amino acid transporter B0+.

DE Amino acid transporter B0+.

ATB0+.

GN Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

OX [1]

SEQUENCE FROM N.A.

RP TISSUE-Mammary gland;

MEDLINE=9376996; PubMed=10446133;

RC Sloan J.L., Mager S.;

"Cloning and functional expression of a human Na(+) and Cl(-)-dependent neutral and cationic amino acid transporter B(0+).";

RT J. Biol. Chem. 274:23740-23745(1999).

EMBL: AF151978; AAD49223.1; -.

RL Genew; HGNC:11047; SLC6A14.

InterPro: IPR00175; Na/ntran_symport.

DR Pfam: PF00209; SNF. 1.

PRINTS: PR00176; NANEUSMPORT.

DR PRODOM: PD000448; Na/ntran_symport; 2.

PROSITE: PS00610; NA_NEUROTAN_SYM_1; 1.

DR PROSITE: PS50267; NA_NEUROTAN_SYM_3; 1.

SEQUENCE 642 AA; 72152 MW; E0FCDD5F173128C0 CRC64;

SO

Query Match

Best Local Similarity

98.2%; Score 1176; DB 4; Length 642;

Matches 213; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy

1 SEETITTTIDLPFKVKKRVPITLGGCLVFLGLGVCYTGAGIYVWHLIDHFCAGWG1 60

Db

427 SIEITTTTIDLPFKVKKRVPITLGGCLVFLGLGVCYTGAGIYVWHLIDHFCAGWG1 486

QY 61 LIAAILELGIIMVYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 120
 DB 487 LIAAILELGIIMVYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 546
 QY 121 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLISCRPASNM 180
 DB 547 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLISCRPASNM 606
 QY 181 GPYLEQHRGERYKDMVVPKKEAGHEIPTVSGSKRPE 216
 DB 607 GPYLEQHRGERYKDMVVPKKEAGHEIPTVSGSKRPE 642

RESULT 2

QY 015003 PRELIMINARY; PRT; 331 AA.
 AC 015003;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE GLYT-1 like (Fragment).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ~~Craetham-D~~ (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RE EMBL: Z96810; CAB09650.1; .
 DR InterPro: IPR000175; Na/ntran_symport.
 DR Pfam: PF00209; SNF; 1.
 DR PRINTS: PR00176; NANEUSMPORT.
 DR PRODOM: PD000448; Na/ntran_symport; 1.
 DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
 FT NON_TER 1
 FT NON_TER 331
 SQ SEQUENCE 331 AA; 37479 MW; 3B761927404F4E57 CRC64;

Query Match 92.4%; Score 1106.5; DB 4; Length 331;
 Best Local Similarity 94.4%; Pred. No. 4.1e-98;
 Matches 204; Conservative 0; Mismatches 3; Indels 9; Gaps 2;

QY 1 SEETITTTIDDLPPKMKKRRVPTIGCCIVLFLGLVCTQAGIYVWHLIDHFCAGWGI 60
 DB 117 SITITTTIDDLPPKMKKRRVPTIGCCIVLFLGLVCTQAGIYVWHLIDHFCAGWGI 175
 QY 61 LIAAILELGIIMVYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 112
 DB 176 LIAAILELGIIMVYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 235
 QY 113 AITFWSVQFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLIS 172
 DB 236 AITFWSVQFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLIS 295
 QY 173 CCRPASNMGPYLEQHRGERYKDMVVPKKEAGHEIPT 208
 DB 296 CCRPASNMGPYLEQHRGERYKDMVVPKKEAGHEIPT 331

RESULT 3

QY 09JMA9 PRELIMINARY; PRT; 638 AA.
 AC 09JMA9;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Colonic system B0+ amino acid transporter.
 GN SLC6A14 OR MCBT0+.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_Taxid=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ady;
 RA Uga S.;
 RT "Mouse Colonic System B0+ Amino Acid Transporter";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB033285; BAA94300.1; .
 DR MGD: MGI:1890216; SLC6a14.
 DR InterPro: IPR000175; Na/ntran_symport.
 DR Pfam: PF00209; SNF; 1.
 DR PRINTS: PR00176; NANEUSMPORT.
 DR PRODOM: PD000448; Na/ntran_symport; 2.
 DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
 DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
 SQ SEQUENCE 638 AA; 71455 MW; 5D5A78E2DFE7612 CRC64;

Query Match 89.6%; Score 1074; DB 11; Length 638;
 Best Local Similarity 86.1%; Pred. No. 1e-94;
 Matches 186; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 SEETITTTIDDLPPKMKKRRVPTIGCCIVLFLGLVCTQAGIYVWHLIDHFCAGWGI 60
 DB 423 SITITTTIDDLPPKMKKRRVPTIGCCIVLFLGLVCTQAGIYVWHLIDHFCAGWGI 482
 QY 61 LIAAILELGIIMVYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 120
 DB 483 LIAAILELGIIMVYGNRFIEDTEMMIGAKRMIFWLMWRACWFVITPILIAIFWLSLV 542
 QY 121 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLISCRPASNM 180
 DB 543 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPIMAIKIIQAKGNIFORLISCRPASNM 602
 QY 181 GPYLEQHRGERYKDMVVPKKEAGHEIPTVSGSKRPE 216
 DB 603 GPYLEQHRGERYKDMVVPKKEAGHEIPTVSGSKRPE 638

RESULT 4

QY 091Y60 PRELIMINARY; PRT; 638 AA.
 AC 091Y60;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Na+ and Cl-coupled neutral and basic amino acid transporter
 DE ATB0+.
 GN SLC6A14 OR ATB0+.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL; TISSUE=COLON;
 RX MEDLINE=21203702; PubMed=11306607;
 RA Hatanaka T., Nakanishi T., Huang W., Leibach F.H., Prasad P.D.,
 RA Ganapathy V., Ganapathy M.E.;
 RT "Na(+) and Cl(-)-coupled active transport of nitric oxide synthase
 RT inhibitors via amino acid transport system B(0,+).";
 RL J. Clin. Invest. 107:1035-1043(2001).
 DR EMBL: AF320226; AAK43541.1; .
 DR MGD: MGI:1890216; SLC6a14.
 DR InterPro: IPR000175; Na/ntran_symport.
 DR Pfam: PF00209; SNF; 1.
 DR PRINTS: PR00176; NANEUSMPORT.
 DR PRODOM: PD000448; Na/ntran_symport; 2.
 DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
 DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
 SQ SEQUENCE 638 AA; 71428 MW; C68188E85F9837 CRC64;

Query Match 89.6%; Score 1074; DB 11; Length 638;
 Best Local Similarity 86.1%; Pred. No. 1e-94;
 Matches 186; Conservative 14; Mismatches 16; Indels 0; Gaps 0;


```

OY 1 SEETITTTTODLPFKVKKRMVPTLGGCLVFLGLGVCYQAGIYVHLLDHFCAAGCI 60
DB 423 SEITTTTTFODLPFKAKRMVPTLGGCLLFLGLGCLVQAGIYVHLLDHFCAAGCI 482
OY 61 LIAAILELVGIWYIGNRFEDTEMIGAKRMIFWLMWRACWFYITPILIAIFWISLV 120
DB 483 LIAAILELVGIWYIGNRFEDTEMIGAKRMIFWLMWRACWFYITPILIAIFWISLV 542
OY 121 QFHRPNYGAIPYPMGVALGCMIVFCIIMIPMAIITIOAKGNIFORLISCRPASNM 180
DB 543 KFHPRVDYADIPYPMGVALGCMIFCIIIMIPMAIITIOAKGNIFORLISCRPASNM 602
OY 181 GPYLEOHNGERYKDMVVPKKEAGHEIPTVSGSRKPE 216
DB 603 GPYLEKHNGERYKDMAPAKETDHEIPTISGSTKPE 638

RESULT 5
O9R183 PRELIMINARY; PRT: 638 AA.
ID 09R183;
AC 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
GN Amino acid transporter B0+.
OS SLC6A14 OR ATB0+.
OC Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Lung;
RX MEDLINE=99376996; PubMed=10446133;
RA Sloan J.L., Mager S.;
RT "Cloning and functional expression of a human Na(+) and Cl(-)-
dependent neutral and cationic amino acid transporter B(0+).";
RL J. Biol. Chem. 274:23740-23745(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Lung;
RA Reveill L., Sloan J.L., Mager S.;
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF161714; AAD49320.1; -.
DR MGD: MGI:1890216; SLC6a14.
DR InterPro: IPR000175; Na/nttran_sympor.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/nttran_sympor. 2.
DR PROSITE: PS00610; NA_NEUTROTAN_SYMP_1; 1.
DR PROSITE: PS50267; NA_NEUTROTAN_SYMP_3; 1.
SQ SEQUENCE 638 AA; 71332 MW; 5607420A394D5092 CRC64;

Query Match 89.6%; Score 1074; DB 11; Length 638;
Best Local Similarity 86.1%; Pred. No. 1e-94;
Matches 186; Conservative 14; Mismatches 16; Indels 0; Gaps 0;
OY 1 SEETITTTTODLPFKVKKRMVPTLGGCLVFLGLGVCYQAGIYVHLLDHFCAAGCI 60
DB 423 SEITTTTTFODLPFKAKRMVPTLGGCLLFLGLGCLVQAGIYVHLLDHFCAAGCI 482
OY 61 LIAAILELVGIWYIGNRFEDTEMIGAKRMIFWLMWRACWFYITPILIAIFWISLV 120
DB 483 LIAAILELVGIWYIGNRFEDTEMIGAKRMIFWLMWRACWFYITPILIAIFWISLV 542
OY 121 QFHRPNYGAIPYPMGVALGCMIVFCIIMIPMAIITIOAKGNIFORLISCRPASNM 180
DB 543 KFHPRVDYADIPYPMGVALGCMIFCIIIMIPMAIITIOAKGNIFORLISCRPASNM 602
OY 181 GPYLEOHNGERYKDMVVPKKEAGHEIPTVSGSRKPE 216
DB 603 GPYLEKHNGERYKDMAPAKETDHEIPTISGSTKPE 638

```

```

RESULT 6
O9D317 PRELIMINARY; PRT: 638 AA.
ID 09D317;
AC 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
GN 9030613J17Rik protein.
OS SLC6A14 OR 9030613J17Rik.
OC Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
Salto T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Salto K.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nakai R., Pesole G., Quackenbush J.,
Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Botfield D., Bolunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Saeki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontaiki S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK018553; BAB31272.1; -.
DR MGD: MGI:1890216; SLC6a14.
DR InterPro: IPR00175; Na/nttran_sympor.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/nttran_sympor. 2.
DR PROSITE: PS00610; NA_NEUTROTAN_SYMP_1; 1.
DR PROSITE: PS50267; NA_NEUTROTAN_SYMP_3; 1.
SQ SEQUENCE 638 AA; 71412 MW; 5D461534187E7612 CRC64;

Query Match 89.0%; Score 1066; DB 11; Length 638;
Best Local Similarity 85.6%; Pred. No. 5.8e-94;
Matches 185; Conservative 14; Mismatches 17; Indels 0; Gaps 0;
OY 1 SEETITTTTODLPFKVKKRMVPTLGGCLVFLGLGVCYQAGIYVHLLDHFCAAGCI 60
DB 423 SEITTTTTFODLPFKAKRMVPTLGGCLLFLGLGCLVQAGIYVHLLDHFCAAGCI 482
OY 61 LIAAILELVGIWYIGNRFEDTEMIGAKRMIFWLMWRACWFYITPILIAIFWISLV 120
DB 483 LIAAILELVGIWYIGNRFEDTEMIGAKRMIFWLMWRACWFYITPILIAIFWISLV 542
OY 121 QFHRPNYGAIPYPMGVALGCMIVFCIIMIPMAIITIOAKGNIFORLISCRPASNM 180
DB 543 KFHPRVDYADIPYPMGVALGCMIFCIIIMIPMAIITIOAKGNIFORLISCRPASNM 602
OY 181 GPYLEOHNGERYKDMVVPKKEAGHEIPTVSGSRKPE 216
DB 603 GPYLEKHNGERYKDMAPAKETDHEIPTISGSTKPE 638

RESULT 7
O9I202 PRELIMINARY; PRT: 791 AA.
ID 09I202

```

AC 091202;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Glycine transporter type 2.
GN SLC6A5 OR GLYT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/c; TISSUE=Brain;
RA Liu Q.-R., Li Q.-F.;
RT "Cloning and expression of mouse glycine transporter 2 (GLYT2).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF411042; AAL17054.1; -
DR MGD; MGI:105090; SLC6A5.
DR InterPro; IPR001920; Asp/Glu_race.
DR InterPro; IPR00175; Na/ntran_symport.
DR Pfam; PF00209; SNF.1.
DR PRINTS; PR00176; NANEUSMPORT.
DR ProDom; PD000448; Na/ntran_symport; 2.
DR PROSITE; PS00923; Asp_Glu_RACEMASE_1; 1.
DR PROSITE; PS00610; NA_NEUROTAN_SYM_1; 1.
DR PROSITE; PS00754; NA_NEUROTAN_SYM_2; 1.
DR PROSITE; PS50267; NA_NEUROTAN_SYM_3; 1.
SQ SEQUENCE 791 AA; 87029 MW; 8AA8B6D0D89EAF CRC64;

Query Match 42.4%; Score 508.5; DB 11; Length 791;
Best Local Similarity 46.9%; Pred. No. 2e-40;
Matches 92; Conservative 32; Mismatches 69; Indels 3; Gaps 1;
QY 3 ETTTITQDLPKRVKMKRVPITIGCCLVLEFLGLVCTVQAGIYWHILIDHFCAGWGI 62
DB 578 ETVTSTISDEPKLRFHKVFTLGCICCFIMGFPMITGGIYMFQVDTYASVALVI 637
QY 63 AALILEVGIIMVGNFIEDTEEMIGAKRMIFWLMRACMFVITPILIAIFWISLYOF 122
DB 638 IALFELVIGISYVGLQRFCEDEIMIGFKNIF--WKVCMAFVTPITLIFILCFSTYQW 694
QY 123 HRPNYGALPYPDWGVAGWCMIVFCIIMIPIMAIKIIOAKGNIFORLISCCRPASWGP 182
DB 695 EPMYGSYRPNMNSVGMILACSVIWMIPIMEVYIKYLAIPRIETRLKIVCSPQPMGP 754
QY 183 YLEOHGGERYKDMVVP 198
DB 755 FLAOhRGERYKNMIDP 770

RESULT 8
O8CFM5 PRELIMINARY; PRT; 799 AA.
AC O8CFM5;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Glycine transporter 2a.
GN GLYT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/c; TISSUE=Brain;
RA Liu Q.-R., Li Q.-F.;
RT "Cloning and expression of mouse sodium-dependent glycine transporter 2 (GLYT2).";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY147186; NANI1408.1; -
SQ SEQUENCE 799 AA; 87891 MW; 47749DIDDDBD3BAB CRC64;

Query Match 42.4%; Score 508.5; DB 11; Length 799;
Best Local Similarity 46.9%; Pred. No. 2e-40;
Matches 92; Conservative 32; Mismatches 69; Indels 3; Gaps 1;
QY 3 ETTTITQDLPKRVKMKRVPITIGCCLVLEFLGLVCTVQAGIYWHILIDHFCAGWGI 62
DB 586 ETVTSTISDEPKLRFHKVFTLGCICCFIMGFPMITGGIYMFQVDTYASVALVI 645
QY 63 AALILEVGIIMVGNFIEDTEEMIGAKRMIFWLMRACMFVITPILIAIFWISLYOF 122
DB 646 IALFELVIGISYVGLQRFCEDEIMIGFKNIF--WKVCMAFVTPITLIFILCFSTYQW 702
QY 123 HRPNYGALPYPDWGVAGWCMIVFCIIMIPIMAIKIIOAKGNIFORLISCCRPASWGP 182
DB 703 EPMYGSYRPNMNSVGMILACSVIWMIPIMEVYIKYLAIPRIETRLKIVCSPQPMGP 762
QY 183 YLEOHGGERYKDMVVP 198
DB 763 FLAOhRGERYKNMIDP 778

RESULT 9
O8VC47 PRELIMINARY; PRT; 633 AA.
AC O8VC47;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Glycine transporter 1.
GN SLC6A9 OR GLYT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021828; AAN21828.1; -
DR MGD; MGI:95760; SLC6A9.
DR InterPro; IPR00175; Na/ntran_symport.
DR Pfam; PF00209; SNF.1.
DR PRINTS; PR00176; NANEUSMPORT.
DR ProDom; PD000448; Na/ntran_symport; 2.
DR PROSITE; PS00610; NA_NEUROTAN_SYM_1; 1.
DR PROSITE; PS00754; NA_NEUROTAN_SYM_2; 1.
DR PROSITE; PS50267; NA_NEUROTAN_SYM_3; 1.
SQ SEQUENCE 633 AA; 70557 MW; 51A0BD971F1B794 CRC64;

Query Match 28.8%; Score 344.5; DB 11; Length 633;
Best Local Similarity 32.4%; Pred. No. 8.6e-25;
Matches 67; Conservative 54; Mismatches 79; Indels 7; Gaps 4;
QY 3 ETTTITQDLPKRVKMKRVPITIGCCLVLEFLGLVCTVQAGIYWHILIDHFCAGWGI 60
DB 405 ETVTALVDEGVGEMWLOK-KTYVTIGVAVAGFLGIPILISQAGIYWLIMDNYASFSL 463
QY 61 LIALILEVGIIMVGNFIEDTEEMIGAKRMIFWLMRACMFVITPILIAIFWISLY 120
DB 464 VVISICMVSIMVYIGHRNFTDIOHMLGPPPLF--FQICRPFVSPALIFILFVY 520
QY 121 QFHRPNYGALPYPDWGVAGWCMIVFCIIMIPIMAIKIIOAKGN-IFORLISCCRPASN 179
DB 521 QYRPITYNHXYGMAVAIFLMAISSVICIPLALFOLCORTGDTLLDRLKATPSPSD 580
QY 180 WGPYLDHGERYKDMVVPKKEAGHEI 206
DB 581 WGPALLEHRTGRYAPITTPSPEDGFEV 607

RESULT 10
O9V7R0

ID Q9Y7R0 PRELIMINARY; PRT; 631 AA.
 AC Q9Y7R0; Q961H9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG8380 protein (CH22929p) (Dopamine transporter).
 GN DAT OR CG8380.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.A., Bouck J., Brockstein P., Brothier P.,
 RA Buttle K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelia S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaisl M., Kalush F., Karpen G.H., Ke Z., Kentonson J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh N.V., McLeod M.P., McPherson D.,
 RA Merkulov G., Mileshina N.V., Mobarry C., Morris A., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferrera S., Frise E., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jaisl M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
 RA Paclob J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Milburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome";
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RP [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunco J., Paclob J., Paragas V., Park S., Phouenavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RN Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Gu H.H., Wu X.;
 RT "Phenylalanine 105 in transmembrane domain 2 of the dopamine
 transporter is involved in cocaine binding.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003806; AAF5986.2;
 DR EMBL: AY051579; AAK93003.1;
 DR EMBL: AF439752; AAI32055.1;
 DR FLYBASE: FBgn004136; DAT.
 DR InterPro: IPR000175; Na/atrian_symport.
 DR PRINTS: PR00176; NANEUSMPORT.
 DR PROSITE: PS00610; NA_NEUROTRAN_SYMPT_1;
 DR PROSITE: PSS0267; NA_NEUROTRAN_SYMPT_3;
 SO SEQUENCE 631 AA; 70330 MW; E7582D1F81E3186 CRC64;
 Query Match 27.7%; Score 332; DB 5; Length 631;
 Best Local Similarity 36.8%; Pred. No. 1,4e-23;
 Matches 64; Conservative 20; Mismatches 77; Indels 4; Gaps 2;
 QY 3 ETITTTTODLEPRKMKMKRYPITIGCCLVLEFLGVCVTQAGTYWHILHFCAGGILTI 62
 DB 427 EAITTAISDEPKI-KNRRELFVAGLPSLVFVGLASCTGGFFHLDRYAGSYILV 485
 QY 63 AAILLEGIIMWYGNFIEDTEKMGAKRMIFLWMRACMFVTPITLLAIFWISLVQF 122
 DB 486 AVFFEAIVASVITGNFSESDIRDMISFPP---GRYQVCNREYVAFLFLFTYGLIGY 542
 QY 123 HRPNYGAIPYDMGVALGMCWIVFCIIIPIMAIKIKIIOAKNIFQRLISCRP 176
 DB 543 EPIITYADYVPSNMANALGMCICASSVYMIPIAIFKLISPGSLRQRFITLTTP 596
 RESULT 11
 ID Q9NB97 PRELIMINARY; PRT; 631 AA.
 AC Q9NB97;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Dopamine transporter.
 GN DAT OR CG8380.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RX MEDLINE-21152910; PubMed-11231083;
 RA Holtschuh J., Ryu S., Aberger F., Drilever W.;
 RT "Dopamine transporter expression distinguishes dopaminergic neurons
 RT from other catecholaminergic neurons in the developing zebrafish
 RT embryo."
 RL Mech. Dev. 101:237-243(2001).
 DR EMBL; AF318177; AAK52449.1; -
 DR ZFIN; ZDB-GENE-010316.1; slc6a3.
 DR InterPro; IPR000175; Na/nttran_symport.
 DR Pfam; PF00209; SNF; 1.
 DR PRINTS; PR00176; NANEUSMPORT.
 DR PRODOM; PD000448; Na/nttran_symport; 1.
 DR PROSITE; PS00610; NA_NEUROTRAN_SYM_1; 1.
 DR PROSITE; PS00754; NA_NEUROTRAN_SYM_2; 1.
 DR PROSITE; PS50267; NA_NEUROTRAN_SYM_3; 1.
 SQ SEQUENCE 629 AA; 70161 MW; EF7CD46C3078C8CB CRC64;

Query Match 27.3%; Score 327; DB 13; Length 629;
 Best Local Similarity 34.9%; Pred. No. 4.1e-23;
 Matches 66; Conservative 33; Mismatches 80; Indels 10; Gaps 3;

OY 3 ETITTTTODLFPPKMKKMRVPTTGGCLVFLGLVCYQAGIYVWHLIDHFCAGWGILI 62
 Db 437 ESVITGLIDEF-KFLHKHRELFTLPIVSTFLSLICVTNGIYVFTLLDHPAAGTSLF 495
 OY 63 AALELVGIITWYGNRFEDTEMMIGAKRMIFLWLRACWFTVTPILLIAIFISLVQF 122
 Db 496 GVLIEAIGVAMFYGVQFSDDIKWMTGGR--NLWRLCKWKLVSPCFLIYVVVSIYTF 552
 OY 123 HRPNTGAIYPDPMGVALGCMIVCIITWIPIMAIKTIQAKNIFORLISCCRP 182
 Db 553 NPKYGSYFFPTWAMWVGCISIMWVPIYATYFKCSLPGSFCDKLAVAITPET---- 608
 OY 183 YLEORGER 191
 Db 609 --DHLVER 615

RESULT 15
 O9ROX6 PRELIMINARY; PRT; 619 AA.
 AC O9ROX6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dopamine transporter.
 GN SLG6A3 OR DAT1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIB/C; TISSUE=Brain;
 RA Brues M., Wieland A., Bonisch H.;
 RT "Molecular cloning and functional expression of the mouse dopamine
 RT transporter."
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ238309; CAB51926.1; -
 DR MGD; MGI:94862; SLG6A3.
 DR InterPro; IPR000175; Na/nttran_symport.
 DR Pfam; PF00209; SNF; 1.
 DR PRINTS; PR00176; NANEUSMPORT.
 DR PRODOM; PD000448; Na/nttran_symport; 1.
 DR PROSITE; PS00610; NA_NEUROTRAN_SYM_1; 1.
 DR PROSITE; PS00754; NA_NEUROTRAN_SYM_2; 1.
 DR PROSITE; PS50267; NA_NEUROTRAN_SYM_3; 1.
 SQ SEQUENCE 619 AA; 68775 MW; F42BA5A959CC3FDE CRC64;

Query Match 26.9%; Score 322; DB 11; Length 619;
 Best Local Similarity 36.2%; Pred. No. 1.2e-22;
 Matches 63; Conservative 32; Mismatches 75; Indels 4; Gaps 2;

OY 3 ETITTTTODLFPPKMKKMRVPTTGGCLVFLGLVCYQAGIYVWHLIDHFCAGWGILI 62
 Db 427 ESVITGLIDEF-QLLHHRRELFTLPIVATFLSLICVANGIYVFTLLDHPAAGTSLF 485
 OY 63 AALELVGIITWYGNRFEDTEMMIGAKRMIFLWLRACWFTVTPILLIAIFISLVQF 122
 Db 486 GVLIEAIGVAMFYGVQFSDDIKWMTGGR--NLWRLCKWKLVSPCFLIYVVVSIYTF 542
 OY 123 HRPNTGAIYPDPMGVALGCMIVCIITWIPIMAIKTIQAKNIFORLISCCRP 176
 Db 543 RPHYGAIVFPDMANALGMIITATSSMAWVPIYATYFKCSLPGSFRELAVAITP 596

Search completed: August 27, 2003, 17:36:25
 Job time : 90 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 09:47:19 ; Search time 378 Seconds
(without alignments)
7812.659 Million cell updates/sec

Title: US-09-940-227-15

Perfect score: 1094

Sequence: 1 gttccagggagctgtacc.....tggttttttaaaaaaaaaa 1094

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1094	100.0	1094	24	ABK81829
2	800.6	73.2	4817	23	ABV25142
3	798	72.9	2098	22	AAH98176
4	641.8	58.7	1929	21	AAH07378
5	633.8	57.9	2177	21	AAH07379
6	280	25.6	382	21	AAH30720
7	152.6	13.9	2394	21	AAH91843
8	152.6	13.9	2394	21	AAH91844

9	149.4	13.7	2394	21	AAH61439
10	149.4	13.7	2394	21	AAH61440
11	149.4	13.7	2394	21	AAH91842
12	149.4	13.7	2394	21	AAH92001
13	149.4	13.7	2397	19	AAV22932
14	149.4	13.7	2397	19	AAV22932
15	147.8	13.5	840	19	AAV22903
16	147.8	13.5	840	19	AAV22904
17	147.8	13.5	2394	21	AAH61441
18	147.8	13.5	2397	19	AAV22931
19	147.8	13.5	2397	19	AAV22915
20	147.8	13.5	2397	19	AAV22916
21	147.8	13.5	2397	19	AAV22917
22	147.8	13.5	2397	19	AAV22918
23	147.8	13.5	2397	19	AAV22920
24	147.8	13.5	2397	19	AAV22921
25	147.8	13.5	2397	19	AAV22924
26	147.8	13.5	2397	19	AAV22925
27	147.8	13.5	2397	19	AAV22926
28	147.8	13.5	2397	19	AAV22927
29	147.8	13.5	2397	19	AAV22928
30	147.8	13.5	2397	19	AAV22929
31	147.8	13.5	2397	19	AAV22930
32	147.8	13.5	2397	19	AAV22910
33	147.8	13.5	2397	19	AAV22911
34	147.8	13.5	2397	19	AAV22912
35	147.8	13.5	2397	19	AAV22913
36	147.8	13.5	2397	19	AAV22914
37	147.8	13.5	2397	19	AAV22905
38	147.8	13.5	2397	19	AAV22909
39	147.8	13.5	2397	19	AAV22906
40	147.8	13.5	2863	21	AAH95536
41	146.2	13.4	2394	21	AAH65003
42	146.2	13.4	2397	19	AAV22919
43	146.2	13.4	2397	19	AAV22923
44	144.6	13.2	589	19	AAV22908
45	139.8	12.8	589	19	AAV22907

ALIGNMENTS

RESULT 1
ABK81829 standard: DNA: 1094 BP.

XX	13-AUG-2002 (first entry)
DT	DNA representing lung specific gene #15.
DE	Lung specific gene; gene therapy; vaccine; lung cancer;
XX	cancer staging; cancer monitoring; cancer diagnosis;
KW	Imaging lung cancer; metastases; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200218576-A2.
PD	
XX	
PD	07-MAR-2002.
XX	
PF	27-AUG-2001, 2001MO-US26684.
XX	
PR	28-AUG-2000, 2000US-228378P.
XX	
PA	(DIAD-) DIADEXUS INC.
XX	
PI	Chen S, Macina RA, Sun Y, Reclon H,
XX	WPI, 2002-434904/46.
DR	
XX	New lung specific genes and their encoded proteins, useful in gene
PT	

inverted

Db 1585 TGGTTTGTATATACGGCTATCCCTTTGATTGCATTTATCTGGTCATTTGGGCATTT 1644
OY 632 CATAGACCTAATTATGGCCAAATTCATACCCCTGATGGGGAGTTCCTTAGCGTGTGT 691
Db 1645 CATAGACCTAATTATGGCCAAATTCATACCCCTGATGGGGAGTTCCTTAGCGTGTGT 1704
OY 692 ATGATGTTTTCGCAATTTATGGATTCATATATGGCATCTATAAATAATTCAGGCT 751
Db 1705 ATGATGTTTTCGCAATTTATGGATTCATATATGGCATCTATAAATAATTCAGGCT 1764
OY 752 AAAGGAACATCTTTCAAGCCCTTATAGTTCGTCAGACCAAGCTTCAACTGGGTCGA 811
Db 1765 AAAGGAACATCTTTCAAGCCCTTATAGTTCGTCAGACCAAGCTTCAACTGGGTCGA 1824
OY 812 TACCTGGAACACATCGTGGGGAAGATATAAGACATGGTAGTCTTAATAAAGAGCT 871
Db 1825 TACCTGGAACACATCGTGGGGAAGATATAAGACATGGTAGTCTTAATAAAGAGCT 1884
OY 872 GGGCATGAATACCTCTGTATGTGGCAGAGAAACCGGAATGAGATCTCATGAAAA 931
Db 1885 GACCATGAATACCTCTGTATGTGGCAGAGAAACCGGAATGAGATCTCATGAAAA 1944
OY 932 AATAATGATTTGATATGATGATTTTATAGAAATAGGGGGAGCCCTTATTTATTTGTGT 991
Db 1945 AATAATGATTTGATATGATGATTTTATAGAAATAGGGGGAGCCCTTATTTATTTGTGT 2004
OY 992 TTAATGATAGGAAATGATACATGATGATGATGATGATGATGATGATGATGATGAT 1051
Db 2005 TTAATGATAGGAAATGATACATGATGATGATGATGATGATGATGATGATGATGAT 2064
OY 1052 AGCAGGATGCAATATAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
Db 2065 AGCAGGATGCAATATAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2098

RESULT 4
AAA07378 standard; DNA; 1929 BP.

AC AAA07378;
XX 30-JUN-2000 (first entry)
DE Human GLYTlike protein coding sequence #1.
XX
KW GLYTlike protein; Glycine transporter protein; human; neuropathic pain;
KW chronic pain; postoperative pain; rheumatoid arthritic pain; neuralgia;
KW neuropathy; algasia; nerve injury; ischaemia; neurodegeneration; stroke;
KW incontinence; inflammatory disorder; spasticity; myoclonus; epilepsy;
KW head trauma; chromosome localisation study; therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200014221-A1.
XX
PD 16-MAR-2000.
XX
PF 03-SEP-1999; 99MO-GH02909.
XX
PR 04-SEP-1998; 98GB-0019405.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Duckworth DM, Evans JR;
XX
XX WPI; 2000-256973/22.
DR P-PSDB; AAY81943.
XX
XX
PT New human glycine transporter, known as GLYTlike, useful in the
PT treatment of diseases such as neuropathic pain, chronic pain,
PT neuralgia, ischaemia, stroke, incontinence, spasticity and epilepsy,
PT comprises a 642 amino acid sequence -

XX
PS Claim 4; page 17; 33pp; English.
XX
CC This sequence encodes a human glycine transporter (GLYT) protein,
CC designated GLYTlike. The GLYTlike polypeptide and the polynucleotide
CC encoding it are useful in the treatment of diseases such as neuropathic
CC pain, pain, chronic pain, postoperative pain, rheumatoid arthritic pain,
CC neuralgia, neuropathies, algasia, nerve injury, ischaemia,
CC neurodegeneration, stroke, incontinence, inflammatory disorders,
CC spasticity, myoclonus, epilepsy and head trauma. The DNA sequence is
CC valuable for chromosome localisation studies. The GLYTlike polypeptide
CC and polynucleotide are also useful in diagnostic assays for detecting
CC diseases associated with in appropriate GLYTlike activity or levels.
XX
SQ Sequence 1929 BP; 521 A; 346 C; 439 G; 623 T; 0 other;

Query Match 58.7%; Score 641.8; DB 21; Length 1929;
Best Local Similarity 99.7%; Pred. No. 4,4e-154;
Matches 643; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 272 GAACAGTACACAAACAAATTCAGATTTATTTCCCAAGCTGATGAAGAAATGAGGCT 331
Db 1285 GAACAGTACACAAACAAATTCAGATTTATTTCCCAAGCTGATGAAGAAATGAGGCT 1344
OY 332 CCCATACCTTTGGGCGTGGCTGTTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391
Db 1345 CCCATACCTTTGGGCGTGGCTGTTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1404
OY 392 GCTGGAATTTACTGGGTTTCATCTGATTTGACCACTCTGCTGCTGCTGCTGCTGCTGCT 451
Db 1405 GCTGGAATTTACTGGGTTTCATCTGATTTGACCACTCTGCTGCTGCTGCTGCTGCTGCT 451
OY 452 GCAGCTATCTGAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 511
Db 1465 GCAGCTATCTGAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 1524
OY 512 GATACAGAAATGATGATTTGGAGCAAGAGTGGATTTGCTGGCTATGTGAGAGCTTGC 571
Db 1525 GATACAGAAATGATGATTTGGAGCAAGAGTGGATTTGCTGGCTATGTGAGAGCTTGC 1584
OY 572 TGGTTTGTATATACGGCTATCCCTTTGATTTGCAATTTATCTGGTCATTTGGTCAATT 631
Db 1585 TGGTTTGTATATACGGCTATCCCTTTGATTTGCAATTTATCTGGTCATTTGGTCAATT 1644
OY 632 CATAGACCTAATTATGGCCAAATTCATACCCCTGATGGGGAGTTCCTTAGCGTGTGT 691
Db 1645 CATAGACCTAATTATGGCCAAATTCATACCCCTGATGGGGAGTTCCTTAGCGTGTGT 1704
OY 692 ATGATGTTTTCGCAATTTATGGATTCATATATGGCATCTATAAATAATTCAGGCT 751
Db 1705 ATGATGTTTTCGCAATTTATGGATTCATATATGGCATCTATAAATAATTCAGGCT 1764
OY 752 AAAGGAACATCTTTCAAGCCCTTATAGTTCGTCAGACCAAGCTTCAACTGGGTCGA 811
Db 1765 AAAGGAACATCTTTCAAGCCCTTATAGTTCGTCAGACCAAGCTTCAACTGGGTCGA 1824
OY 812 TACCTGGAACACATCGTGGGGAAGATATAAGACATGGTAGTCTTAATAAAGAGCT 871
Db 1825 TACCTGGAACACATCGTGGGGAAGATATAAGACATGGTAGTCTTAATAAAGAGCT 1884
OY 872 GGGCATGAATACCTCTGTATGTGGCAGAGAAACCGGAATGCA 916
Db 1885 GGGCATGAATACCTCTGTATGTGGCAGAGAAACCGGAATGCA 1929

RESULT 5
AAA07379 standard; DNA; 2177 BP.

AC AAA07379;
XX
XX
XX 30-JUN-2000 (first entry)
XX

DE		Human GLYTLIKE protein coding sequence #2.
XX		
KW	GLYTlike protein; Glycine transporter protein; human; neuropathic pain;	
KM	chronic pain; postoperative pain; rheumatoid arthritic pain; neuralgia;	
KV	neuropathy; algesia; nerve injury; ischemia; neurodegeneration; stroke;	
KW	incontinence; inflammatory disorder; spasticity; myoclonus; epilepsy;	
KX	head trauma; chromosome localisation study; therapy; ss.	
OS	Homo sapiens.	
PN	WO200014221-A1.	
PD	16-MAR-2000.	
PP	03-SEP-1999; 99WO-GB02909.	
PR	04-SEP-1998; 98GB-0019405.	
PA	(SMIK) SMITHKLINE BEECHAM PLC.	
PI	Duckworth DM, Evans JR:	
DR	WPI: 2000-256973/22.	
PS	P-PsDB: AAYB1944.	
PT	New human glycine transporter, known as GLYTlike, useful in the	
PT	treatment of diseases such as neuropathic pain, chronic pain,	
PT	neuralgia, ischemia, stroke, incontinence, spasticity and epilepsy,	
PT	comprises a 642 amino acid sequence -	
XX		
XX	Claim 11: Page 18; 33pp: English.	
XX		
CC	This sequence encodes a human glycine transporter (GlyT) protein,	
CC	designated GLYTlike. The GLYTlike polypeptide and the polynucleotide	
CC	encoding it are useful in the treatment of diseases such as neuropathic	
CC	pain, pain, chronic pain, postoperative pain, rheumatoid arthritic pain,	
CC	neuralgia, neuropathies, algesia, nerve injury, ischemia,	
CC	neurodegeneration, stroke, incontinence, inflammatory disorders,	
CC	spasticity, myoclonus, epilepsy and head trauma. The DNA sequence is	
CC	valuable for chromosome localisation studies. The GLYTlike polypeptide	
CC	and polynucleotide are also useful in diagnostic assays for detecting	
CC	diseases associated with in appropriate GLYTlike activity or levels.	
XX		
SQ	Sequence 2177 BP; 581 A; 408 C; 520 G; 668 T; 0 other;	
	Query Match 57.9%; Score 633.8; DB 21; Length 2177;	
	Best Local Similarity 98.9%; Pred. No. 5.1e-152;	
	Matches 638; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
OY	272 GAACGATCACAACAGCAATTCAAGATTTATTCCTCAAAGTAGAAGAAATGAGGGTT	331
Db	1533 GAACGATCACAACAGCAATTCAAGATTTATTCCTCAAAGTAGAAGAAATGAGGGTT	1592
OY	332 CCCATAACTTTGGGCGCTGCTGTGGTTTTGTTCTCCTGGTGCGTGTGTGACTCAG	391
Db	1593 CCCATAACTTTGGGCGCTGCTGTGGTTTTGTTCTCCTGGTGCGTGTGTGACTCAG	1652
OY	392 GCCTGAATTTACGTGGGTTCACTCATGATGACCACCTTGCTGTGATGGGCATTTTAAT	451
Db	1653 GCCTGAATTTACGTGGGTTCACTCATGATGACCACCTTGCTGTGATGGGCATTTTAAT	1712
OY	452 GCACCTTACTGAGCTAGTTGGATCATCGATTCGATTTATGGAGGAGACAGATTCATTGAG	511
Db	1713 GCACCTTACTGAGCTAGTTGGATCATCGATTCGATTTATGGAGGAGACAGATTCATTGAG	1772
OY	512 GATACAGAATGATGATTTGGAGCAAGAGGTGATATTCTGGCTATGSGTGAGAGCTTGC	571
Db	1773 GATACAGAATGATGATTTGGAGCAAGAGGTGATATTCTGGCTATGSGTGAGAGCTTGC	1832
OY	572 TGCTTTGTAATTTACGCCATCCCTTTGATTTGCATATTTATTCGTGCAATGGTGCATATT	631
Db	1833 TGCTTTGTAATTTACGCCATCCCTTTGATTTGCATATTTATTCGTGCAATGGTGCATATT	1892

QY	632	CATACACCTAAATTTATGGGGCAATTCCTCATTACCCCTGACGTGGGGAGTTGCTTTAGCGCTG	691		
Db	1893	CATAGACCTAAATTTTGGGGCAATTCCTCATTACCCCTGACGTGGGGAGTTGCTTTAGCGCTG	1955		
QY	692	ATGATTTGTTTCTGCAATTAATTTGGATTCCAATTAATGCTATCATTAATAATTAATCAGGCT	751		
Db	1953	ATGATTTGTTTCTGCAATTAATTTGGATTCCAATTAATGCTATCATTAATAATTAATCAGGCT	2012		
QY	752	AAAGAAACATCTTTCAACGCCCTTAATAGTTGCTGCAGACACAGCTTCTAACTGGGCTCCA	811		
Db	2013	AAAGAAACATCTTTCAACGCCCTTAATAGTTGCTGCAGACACAGCTTCTAACTGGGCTCCA	2072		
QY	812	TACCTGGAACAACATCGTGGGGAAGATTAATGAAGCATGATGTTCCATAAAAAAGAGGCT	871		
Db	2073	TACCTGGAACAACATCGTGGGGAAGATTAATGAAGCATGATGTTCCATAAAAAAGAGGCT	2133		
QY	872	GGCCATGAATTAACCTACTGTTAGTGGCAGCAGAAAAACCGGAAATGA	916		
Db	2133	GACCATGAATTAACCTACTGTTAGTGGCAGCAGAAAAACCGGAAATGA	2177		
RESULT 6					
AAH30720	AAH30720 standard; cDNA; 382 BP.				
AC	AAH30720;				
XX	27-JUL-2001 (first entry)				
DT					
XX	Human colon cancer cell line Kml2L4-A cDNA library derived sequence #654.				
DE					
XX					
KM	Human: diagnosis; colon cancer; cancer; malignant; chromosome mapping;				
XX	detection; colon cancer cell line Kml2L4-A; ss.				
OS	Homo sapiens.				
XX					
PN	WO200018916-A2.				
PD	06-APR-2000.				
PF					
XX	23-SEP-1999;	99WO-US22226.			
PR	28-SEP-1998;	98US-0102161.			
PR	28-SEP-1998;	98US-0102180.			
PR	29-SEP-1998;	98US-0102380.			
PR	08-OCT-1998;	98US-0103815.			
PR	27-OCT-1998;	98US-0105877.			
PA	(CHIR) CHIRON CORP.				
PA	(HYSE-) HYSEQ INC.				
PI	Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;				
PI	Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;				
PI	Lamson G, Derman R, Cirkenjakov R, Dickson M, Derman S, Labat I;				
PI	Leshkowitz D, Kita D, Garcia V, Jones LW, Staechelin B;				
XX	WPI: 2000-293155/25.				
XX	Polynucleotide library comprising 1079 defined sequences, useful in				
PT	the form of an array to detect cancer or susceptibility to cancer -				
XX	Claim 1; Page 372; 502pp; English.				
XX					
CC	The present invention describes a library of polynucleotides comprising				
CC	1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described				
CC	are: (1) an isolated polynucleotide (I) having at least 90% identity to				
CC	one of the 1079 sequences; (2) a recombinant host cell containing (I);				
CC	(3) an isolated polypeptide (II) encoded by (1); (4) an antibody that				
CC	specifically binds to (II); (5) a vector comprising (I); and (6) a method				
CC	of detecting differentially expressed genes correlated with a cancerous				
CC	state of a mammalian cell comprising detecting a gene product encoded by				
CC	65 of the 1079 sequences given in the specification. The polynucleotides				
CC	are used to monitor patients having (or susceptible) to cancer to detect				

CC potentially malignant events at a molecular level before they are
CC detectable at a gross morphological level. The polynucleotides are also
CC useful for monitoring the efficacy of various therapies and preventive
CC interventions. Polynucleotide probes based on the disclosed sequences
CC are useful for chromosome mapping and detection of transcription levels.
CC The 1079 polynucleotide sequences were derived from a human colon cancer
CC cell line Kml214-A cDNA library.
XX
SQ Sequence 382 BP; 89 A; 68 C; 92 G; 132 T; 1 other;

Query Match 25.6%; Score 280; DB 21; Length 382;
Best local Similarity 99.6%; Pred. No. 8.8e-62;
Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 272 GAAGAGTACACACAAACATTTATTTCCCAAGTGATGAAGAAATGAGGGTT 331
DB 96 GAACAGTACACACAAACATTTATTTCCCAAGTGATGAAGAAATGAGGGTT 155
QY 332 CCCATTAACCTTGGGCGTGGTGTGTTTCTCCCTGGTCTGTGTGACTCAG 391
DB 156 CCCATTAACCTTGGGCGTGGTGTGTTTCTCCCTGGTCTGTGTGACTCAG 215
QY 392 GCTGAATTTACTGGGTCATCTGATTTGACCACTTCTGCTGGATGGGCATTTAAT 451
DB 216 GCTGAATTTACTGGGTCATCTGATTTGACCACTTCTGCTGGATGGGCATTTAAT 275
QY 452 GCAGCATACCTGAGCTACTTGAATCATCTGATTTAGAGGGAACAGATTCAATGAG 511
DB 276 GCAGCATACCTGAGCTACTTGAATCATCTGATTTAGAGGGAACAGATTCAATGAG 335
QY 512 GATACGAAATGATGATTTGAGCAGCAAGAGGTGATTTCTG 552
DB 336 GATACGAAATGATGATTTGAGCAGCAAGAGGTGATTTCTG 376

RESULT 7
AAA91843
ID AAA91843 standard; cDNA; 2394 BP.
XX
AC AAA91843;
DT 23-FEB-2001 (first entry)
XX
DE Human glycine transporter type 2 protein mutant coding sequence #3.
XX
KW Human glycine transporter type 2; GLYT2; muscle spasticity; myoclonus;
KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;
KW dystonia; Huntington's disease; amyotrophic lateral sclerosis;
KW mutant; muten; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..2394
FT /tag- a
FT /Product- "mutant Glyt2 protein"
FT misc_difference 70
FT /tag- b
FT /note- "wild-type G substituted by A"
FT misc_difference 77
FT /tag- c
FT /note- "wild-type C substituted by T"
FT misc_difference 220
FT /tag- d
FT /note- "wild-type C substituted by T"
FT misc_difference 244
FT /tag- e
FT /note- "wild-type C substituted by T"
FT misc_difference 266
FT /tag- f
FT /note- "wild-type C substituted by T"
FT misc_difference 304

FT /tag- g
FT /note- "wild-type A substituted by G"
FT misc_difference 463
FT /tag- h
FT /note- "wild-type A substituted by G"
FT misc_difference 521
FT /tag- i
FT /note- "wild-type T substituted by A"
FT misc_difference 562
FT /tag- j
FT /note- "wild-type A substituted by G"
FT misc_difference 583
FT /tag- k
FT /note- "wild-type T substituted by C"
FT misc_difference 596
FT /tag- l
FT /note- "wild-type A substituted by G"
FT misc_difference 678
FT /tag- m
FT /note- "wild-type A substituted by G"
FT misc_difference 681
FT /tag- n
FT /note- "wild-type T substituted by C"
FT misc_difference 745
FT /tag- o
FT /note- "wild-type G substituted by T"
FT misc_difference 750
FT /tag- p
FT /note- "wild-type G substituted by C"
FT misc_difference 765
FT /tag- q
FT /note- "wild-type C substituted by T"
FT misc_difference 777
FT /tag- r
FT /note- "wild-type A optionally substituted by C"
FT misc_difference 917
FT /tag- s
FT /note- "wild-type T substituted by C"
FT misc_difference 1085
FT /tag- t
FT /note- "wild-type A substituted by T"
FT misc_difference 1256
FT /tag- u
FT /note- "wild-type G substituted by A"
FT misc_difference 1292
FT /tag- v
FT /note- "wild-type T substituted by C"
FT misc_difference 1299
FT /tag- w
FT /note- "wild-type C substituted by A"
FT misc_difference 1325
FT /tag- x
FT /note- "wild-type C substituted by A"
FT misc_difference 1364
FT /tag- y
FT /note- "wild-type C substituted by A"
FT misc_difference 1374
FT /tag- z
FT /note- "wild-type G substituted by C"
FT misc_difference 1392
FT /tag- aa
FT /note- "wild-type C substituted by A"
FT misc_difference 1454
FT /tag- ab
FT /note- "wild-type T substituted by C"
FT misc_difference 1478
FT /tag- ac
FT /note- "wild-type A substituted by G"
FT misc_difference 1617
FT /tag- ad
FT /note- "wild-type A substituted by T"
FT misc_difference 1744
FT /tag- ae

FT /+tag- k
/note- "wild-type T substituted by C"
FT misc_difference 745
FT /+tag- l
/note- "wild-type G substituted by T"
FT misc_difference 750
FT /+tag- m
/note- "wild-type G substituted by C"
FT misc_difference 765
FT /+tag- n
/note- "wild-type C substituted by T"
FT misc_difference 777
FT /+tag- o
/note- "wild-type G substituted by C"
FT misc_difference 867
FT /+tag- p
/note- "wild-type A substituted by G"
FT misc_difference 917
FT /+tag- q
/note- "wild-type T substituted by C"
FT misc_difference 1256
FT /+tag- r
/note- "wild-type G substituted by A"
FT misc_difference 1292
FT /+tag- s
/note- "wild-type T substituted by C"
FT misc_difference 1325
FT /+tag- t
/note- "wild-type C substituted by A"
FT misc_difference 1364
FT /+tag- u
/note- "wild-type C substituted by A"
FT misc_difference 1374
FT /+tag- v
/note- "wild-type G substituted by C"
FT misc_difference 1392
FT /+tag- w
/note- "wild-type C substituted by A"
FT misc_difference 1454
FT /+tag- x
/note- "wild-type T substituted by C"
FT misc_difference 1478
FT /+tag- y
/note- "wild-type A substituted by G"
FT misc_difference 1854
FT /+tag- z
/note- "wild-type T substituted by C"
FT misc_difference 1949
FT /+tag- aa
/note- "wild-type T substituted by A"
FT misc_difference 1959
FT /+tag- ab
/note- "wild-type T substituted by C"
FT misc_difference 2130
FT /+tag- ac
/note- "wild-type T substituted by C"
XX W0200029564-A2.
XX PD 25-MAY-2000.
XX PF 09-NOV-1999; 99MO-US26534.
XX PR 12-NOV-1998; 98US-0191468.
XX (GLIA-) GLIATECH INC.
XX Gallagher MJ, Burgess LH, Brunden KR;
XX WPI; 2000-38780/33.
XX New nucleic acid encoding a human glycine transporter type 2 (GLYT2),
PT useful for expressing GLYT2 which can then be used for detecting

PT whether an animal has autoimmune antibodies against a glycine
PT transporter -
XX Claim 52; Page -, 175pp; English.
XX
XX The present sequence is a mutant version of the coding sequence of the
CC human glycine transporter type 2 (GLYT2). Glycine is involved in
CC neurotransmission in the spinal cord and cerebellum, where glycine
CC receptor inhibition results in pain transmission. The GLYT2 gene and
CC protein could, therefore, be used to identify inhibitory compounds which
CC prevent GLYT2 activity, causing a decrease in muscle hyperactivity. They
CC could be used to treat diseases or conditions associated with muscle
CC contraction, including muscle spasticity (such as that due to epilepsy,
CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal
CC cord injury and dystonia), myoclonus, epilepsy, Huntington's disease,
CC amyotrophic lateral sclerosis and pain.
CC Note: This sequence is not shown in the specification, but is derived
CC from the hGLYT2 wild type sequence shown in SEQ ID NO: 121 (AAA61440).
XX
SQ Sequence 2394 BP; 508 A; 663 C; 628 G; 595 T; 0 other;

Query Match 13.9%; Score 152.6; DB 21; Length 2394;
Best Local Similarity 55.0%; Pred. No. 6,3e-29;
Matches 327; Conservative 0; Mismatches 259; Indels 9; Gaps 1;

QY 264 CCTCTGAGAGAACGACACACACATTCAGATTATTTCCCAAGTGAAGAA 323
DB 1742 CCTCCATCGAGACAGATAGACCTCCATCTCAGACAGTTCCAGTACGACAC 1801
QY 324 TGAGGGTCCCATACATTTGGGCTGCTGCTGTTGTTCTCTCTGCTGCTG 383
DB 1802 ACAAGCCAGATGTTACTGCGGCTGCTGATTTGTTCTTCATCAGGGTTTCCCAATGA 1861
QY 384 TGACATGAGCTGGAATTTACTGCTCATCTGATTTACACATTCCTGTGTGATGGGCA 443
DB 1862 TCACCTGAGGTGAATTTACATGTTTACCTGCTGTGACACGATGCTGCTCTATGCC 1921
QY 444 TTTTAAATTCAGCTATACGAGCTAGTGAATTCATGATTTATGAGGAGAACAGAT 503
DB 1922 TTGTCATCATTTGCAATTTTGAAGCTGAGAGGATCTCCATGATGCTTGCAGAAAT 1981
QY 504 TCATTGAGATACAGAAATGATGATTTGAGCAAAAGGTGATTTCTGCTATGCTGA 563
DB 1982 TCTGTGAAGATATGAGATGATGATTTGATTCAGCCTAACATCTTC-----TGA 2032
QY 564 GAGCTTGCTGTTTGAATTTAGCGCTATCCTTTGATTCAGATTTATCTGCTATG 623
DB 2033 AAGTCTGCTGGCATTTGTATACCCCAACATTTTAACTTTATCTTTGCTTCACTTT 2092
QY 624 TGCATTTGATAGCTAATTTATGAGCGCAATTCATACCTGATGAGGAGTTGCTTTAG 683
DB 2093 ACCAGTGGAGAACCAATGACCTATGCTTTAGCGCTAACCTATGCTCATGCTGCTCG 2152
QY 684 GCTGTGATGATGTTTCTGCTATTTTGAATTTGATTTGCTATCATTAATAATA 743
DB 2153 GATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2212
QY 744 TTGAGGCTAAGGAACATCTTCAACGCTTATAGTTGCTGACAGACAGCTTCAACT 803
DB 2213 ATCTGCCCCCTGAGAAATTTATGAGAGCTGAAGTTGCTGCTGCTGCTGCTGCTGCT 2272
QY 804 GGGGTCCATACCTGGAACAACATCTGAGGAAAGATTAAGACATGATGATTC 858
DB 2273 GGGGCCCATTTCTAGCTCAACACCGGGGAGGCTTACAAAGACATGATGACCC 2327

RESULT 9
AAA61439
ID AAA61439 standard; cDNA; 2394 BP.
XX AC AAA61439;
XX DT 23-FEB-2001 (first entry)

XX DE Human glycine transporter type 2 protein coding sequence #1.
 XX KW Human glycine transporter type 2; GlyT2; muscle spasticity; myoclonus;
 KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;
 KW dystonia; Huntington's disease; amyotrophic lateral sclerosis; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..2394
 XX FT /*tag= a
 XX FT /product= "GlyT2"
 XX PN W0200029564-A2.
 XX PD 25-MAY-2000.
 XX PE 09-NOV-1999; 99MO-US26534.
 XX PR 12-NOV-1998; 98US-0191468.
 XX PA (GLIA-) GLIATECH INC.
 XX PI Gallagher MJ, Burgess LH, Brunden KR;
 XX DR MPI; 2000-387780/33.
 XX DR P-PSDB; AAB09896.
 XX PT New nucleic acid encoding a human glycine transporter type 2 (GlyT2),
 PT useful for expressing GlyT2 which can then be used for detecting
 PT whether an animal has autoimmune antibodies against a glycine
 PT transporter -
 XX PS Claim 14; Page 157-160; 175pp; English.
 CC The present sequence is one version of the coding sequence of the
 CC human glycine transporter type 2 (GlyT2). Glycine is involved in
 CC neurotransmission in the spinal cord and cerebellum, where glycine
 CC receptor inhibition results in pain transmission. The GlyT2 gene and
 CC protein could, therefore, be used to identify inhibitory compounds which
 CC prevent GlyT2 activity, causing a decrease in muscle hyperactivity. They
 CC could be used to treat diseases or conditions associated with muscle
 CC contraction, including muscle spasticity (such as that due to epilepsy,
 CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal cord
 CC injury and dystonia), myoclonus, epilepsy, Huntington's disease,
 CC amyotrophic lateral sclerosis and pain.
 CC SO Sequence 2394 BP; 505 A; 659 C; 630 G; 600 T; 0 other;
 Query Match 13.7%; Score 149.4; DB 21; Length 2394;
 Best Local Similarity 54.6%; Pred. No. 4.2e-28;
 Matches 325; Conservative 0; Mismatches 261; Indels 9; Gaps 1;
 QY 264 CCTCTGAAGAAAGCATGACACAAACATTCAGATTATTTCCCAAGATGAGAGAAA 323
 DB 1742 CCTCATGACAGACAGACAGCTCCATCTCAGACAGATTCCACAGTACGACAC 1801
 QY 324 TGAGGGTTCATTAACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383
 DB 1802 ACAAGCCAGCTGTTTACTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1861
 QY 384 TGACTCAGGCTGGAATTTACTGGTTCATCTGATTGACCACTTCTGCTGATGGGCA 443
 DB 1862 TCACCTCAGGCTGGAATTTACTGATGTTTACATGTTGACGACATATGCTCTCCTA 1921
 QY 444 TTTTATTTGACCTTATACAGAGCTTGTGGAATCATCTGGAATTTATGAGGGAACAGAT 503
 DB 1922 TTGTCTCATTTGCTCATTTTGTGAGCTGCTGGGATCTCTTATGTATGCTTGCAGAA 1981
 QY 504 TCATTGAGGATACAGAAATGATGATGAGCAAAAGAGTGATATCTGCTATGATGATG 563
 DB 1982 TCTGTGAAGATATAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2032
 1982 TCTGTGAAGATATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

QY 564 GAGCTTCTGCTGTTTGAATTAACGCTATTCCTTTGATTCGAATATTATTCGTCATTTGG 623
 DB 2033 AAGCTGCTGCTGGGATTTGTTATGTAACCCCAACCTTTATACCTTTATTCGTCATTTT 2092
 QY 624 TGCAATTTCTATAGACCTTAATTAAGGCGCAATTCATACCTGAGCTGGAGTTCCTTAG 683
 DB 2093 ACCAGTGGGAACCCATGACCTATGAGCTTTTACCGTATCTTAACCTGATGCTGCTG 2152
 QY 684 GCTGTGTATGATGATGATTTTCTGCAATTTATGATTCATTAATGATGATGATGATGAT 743
 DB 2153 GATGGCTAATGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212
 QY 744 TTCAGGCTTAAGAAACATCTTCAACGCTTATTAAGTGTCTGCAACGCTTCTAACT 803
 DB 2213 ATCTGGCCCTGGAAGATTTATGACAGGCTGAAGTGTGTGCTGCTGCTGCTGCTG 2272
 QY 804 GGGGTCCATFACCTGGAACAAATCGTGGGGAAGATATTAAGACATGATGATGCC 858
 DB 2273 GGGGCCATTTCTTACTCAACACCGGAGGAGCTTACAAAGAACATGATGATGCC 2327
 RESULT 10
 ID AAA61440 standard; cDNA; 2394 BP.
 XX AAA61440;
 AC AAA61440;
 XX 23-FEB-2001 (first entry)
 DT Human glycine transporter type 2 protein consensus coding sequence.
 XX DE Human glycine transporter type 2; GlyT2; muscle spasticity; myoclonus;
 XX KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;
 XX KW dystonia; Huntington's disease; amyotrophic lateral sclerosis; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..2394
 XX FT /*tag= a
 XX FT /product= "GlyT2"
 XX PN W0200029564-A2.
 XX PD 25-MAY-2000.
 XX PE 09-NOV-1999; 99MO-US26534.
 XX PR 12-NOV-1998; 98US-0191468.
 XX PA (GLIA-) GLIATECH INC.
 XX PI Gallagher MJ, Burgess LH, Brunden KR;
 XX DR MPI; 2000-387780/33.
 XX DR P-PSDB; AAB09897.
 XX PT New nucleic acid encoding a human glycine transporter type 2 (GlyT2),
 XX PT useful for expressing GlyT2 which can then be used for detecting
 XX PT whether an animal has autoimmune antibodies against a glycine
 XX PT transporter -
 XX PS Claim 13; Page 163-166; 175pp; English.
 CC The present sequence is the consensus coding sequence of the human
 CC glycine transporter type 2 (GlyT2). The researchers sequenced a number of
 CC clones from spinal cord and astrocytoma cells to produce this consensus
 CC version. Glycine is involved in neurotransmission in the spinal cord and
 CC cerebellum, where glycine receptor inhibition results in pain
 CC transmission. The GlyT2 gene and protein could, therefore, be used to
 CC identify inhibitory compounds which prevent GlyT2 activity, causing a
 CC decrease in muscle hyperactivity. They could be used to treat diseases or


```

Db 1742 CCTCCATCGAGACCATAGTACCTCCATCTCAGACGAGATTCCCAAGTACCTAGCACAC 1801
Qy 324 TGAGGGTTCCCATTAAGTTGGGCTGGCTGGTTGTTCTCTCTGCTCTGCTG 383
Db 1802 ACAAGCCAGTGTCTAGCTGGGCTGGCTGATTTGTTCTTCATCATGGGTTTCCAAATGA 1861
Qy 384 TGACTCAGGCTGAATTTACTGGGTTCTATCTGATTTGACCACTTGTCTGGATGGGCA 443
Db 1862 TCACCTCAGGGTGAATTTACTATGTTTCAGCTGTGGACACCTATGCTCCCTCATATGCC 1921
Qy 444 TTTTAATTCAGCTATCTGAGCTAGTTGGAATCATCTGATTTATGAGGAGAAAGAT 503
Db 1922 TTGTCATCATCTTTCATTTTGAGCTCGTGGGATCTCTATATGTTATAGGCTTGCAAAAGAT 1981
Qy 504 TCATTGAGATACAGAAATGATGATTTGAGCAAGAGTGGATTTCTGCTATAGTGA 563
Db 1982 TCTGTGAAGATATGAGATGATGATTTGATTTCCAGCTTCAACATCTTC-----TGGA 2032
Qy 564 GAGCTTCTGCTGTTTGAATTAAGGCTATTCCTTTTGAATTCATTTATCTGTCATTTGG 623
Db 2032 AAGTCTGCTGGGCAATTTGTAAACCCCAACCAATTTTAACCTTTATCTTGTTCAGCTTTT 2092
Qy 624 TGCATTTTCATTAAGCAATTAATGAGCGCAATTCATACCTGAGTGGGAGTTCCTTTAG 683
Db 2092 ACCGCTGAGGAGCCCATGACCTATGAGCTCTTAACCCCTTCACTGATGCTCATGCTGCTCG 2152
Qy 684 GCTGCTGTATGATTTGTTCTGCAATTAATTTGATTTCAATTAATGATTAATTAATAAATA 743
Db 2152 GATGGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212
Qy 744 TTTAGGCTTAAGCAACATCTTTCAACGCTTATTAAGTCTGTCGACAGCAGCTTCTACT 803
Db 2212 ATCTGGGCTCTGGAAGATTTATGAGAGGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2272
Qy 804 GGGGTCATACCTGAGAACATGTCGGGAAAGATTAATAAGCATGATGTTGCTTC 858
Db 2272 GGGGCCATTTCTTATGCTCAACACCGCGGGAGCCTTACAAAGACATGATGACCC 2327

RESULT 12
AAA92001
ID AAA92001 standard; cDNA; 2394 BP.
XX
AC AAA92001:
XX
DT 23-FEB-2001 (first entry)
XX
DE Human glycine transporter type 2 protein mutant coding sequence #5.
XX
KW Human glycine transporter type 2; GlyT2; muscle spasmodic; myoclonus;
KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;
KW dystonia; Huntington's disease; amyotrophic lateral sclerosis;
KW mutant; mutelin; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..2394
FT CDS
FT /note- "mutant GlyT2 protein"
FT /tag- a
FT /product- "mutant GlyT2 protein"
FT /tag- b
FT /note- "wild-type C substituted by T"
FT /tag- c
FT /note- "wild-type A substituted by G"
FT /tag- d
FT /note- "wild-type T substituted by C"
FT /tag- e
FT /note- "wild-type T substituted by C"
PN
XX
XX MO200029564-A2.

```

```

PD 25-MAY-2000.
XX
XX 09-NOV-1999; 99WO-0526334.
XX
PR 12-NOV-1998; 98US-0191468.
XX
PA (GLIA-) GLINTECH INC.
XX
PI Gallagher MJ, Burgess LH, Brunden KR;
XX
XX WPI: 2000-387780/33.
XX
XX New nucleic acid encoding a human glycine transporter type 2 (GlyT2),
XX useful for expressing GlyT2 which can then be used for detecting
XX whether an animal has autoimmune antibodies against a glycine
XX transporter.
XX
XX Claim 53; Page -: 175pp; English.
XX
XX The present sequence is a mutant version of the coding sequence of the
XX human glycine transporter type 2 (GlyT2). Glycine is involved in
XX neurotransmission in the spinal cord and cerebellum, where glycine
XX receptor inhibition results in pain transmission. The GlyT2 gene and
XX protein could, therefore, be used to identify inhibitory compounds which
XX prevent GlyT2 activity, causing a decrease in muscle hyperactivity. They
XX could be used to treat diseases or conditions associated with muscle
XX contraction, including muscle spasticity (such as that due to epilepsy,
XX stroke, head trauma, multiple sclerosis, neuronal cell death, spinal
XX cord injury and dystonia), myoclonus, Huntington's disease,
XX amyotrophic lateral sclerosis, epilepsy, Huntington's disease,
XX CC Note: This sequence is not shown in the specification, but is derived
XX from the hGlyT2 wild type sequence shown in SEQ ID NO: 121 (AAA61440).
XX
SO Sequence 2394 BP; 506 A; 660 C; 629 G; 599 T; 0 other:

Query Match 13.7%; Score 149.4; DB 21; Length 2394;
Best Local Similarity 54.6%; Pred. No. 4.2e-28;
Matches 325; Conservative 0; Mismatches 261; Indels 9; Gaps 1;

Qy 264 CCTGTGAAGAAAGATGATCAACAACATTTCAATTTATTTCCCAAGTATGAAAGAAA 323
Db 1742 CCTCCATCGAGACCATAGTACCTCCATCTCAGACGAGATTCCCAAGTACCTAGCACAC 1801
Qy 324 TGAGGGTTCCCATTAAGTTGGGCTGGCTGGTTGTTCTCTCTGCTGCTGCTGCTGCTG 383
Db 1802 ACAAGCCAGTGTCTAGCTGGGCTGGCTGATTTGTTCTTCATCATGGGTTTCCAAATGA 1861
Qy 384 TGACTCAGGCTGAATTTACTGGGTTCTATCTGATTTGACCACTTGTCTGGATGGGCA 443
Db 1862 TCACCTCAGGGTGAATTTACTATGTTTCAGCTGTGGACACCTATGCTCCCTCATATGCC 1921
Qy 444 TTTTAATTCAGCTATCTGAGCTAGTTGGAATCATCTGATTTATGAGGAGAAAGAT 503
Db 1922 TTGTCATCATCTTTCATTTTGAGCTCGTGGGATCTCTATATGTTATAGGCTTGCAAAAGAT 1981
Qy 504 TCATTGAGATACAGAAATGATGATTTGAGCAAGAGTGGATTTCTGCTATAGTGA 563
Db 1982 TCTGTGAAGATATGAGATGATGATTTGATTTCCAGCTTCAACATCTTC-----TGGA 2032
Qy 564 GAGCTTCTGCTGTTTGAATTAAGGCTATTCCTTTTGAATTCATTTATCTGTCATTTGG 623
Db 2032 AAGTCTGCTGGGCAATTTGTAAACCCCAACCAATTTTAACCTTTATCTTGTTCAGCTTTT 2092
Qy 624 TGCATTTTCATTAAGCAATTAATGAGCGCAATTCATACCTGAGTGGGAGTTCCTTTAG 683
Db 2092 ACCGCTGAGGAGCCCATGACCTATGAGCTCTTAACCCCTTCACTGATGCTCATGCTGCTCG 2152
Qy 684 GCTGCTGTATGATTTGTTCTGCAATTAATTTGATTTCAATTAATGATTAATTAATAAATA 743
Db 2152 GATGGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2212
Qy 744 TTTAGGCTTAAGCAACATCTTTCAACGCTTATTAAGTCTGTCGACAGCAGCTTCTACT 803

```

Db 2273 ATCTGCGCCCTGGAGAAFTTTATTAGAGAGCCTGAAGTGTGGTGTCGCCACAGCCGAGACT 2272

OY 804 GGGGTCATACCTTGGAACAACATCTGTGGGGAAGAATATATAAGACATGTAAGTC 858
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2273 GGCGCCCATCTTAGCTCAACACCGCGGGAGCGTTACMAAGACATGATGCACC 2327

RESULT 13
ID AAV22922 standard; cDNA; 2397 BP.
XX AAV22922;
AC
XX
DE 29-JUL-1998 (first entry)
XX
Human GLYT-2 transporter T2198C mutant gene sequence.
XX
Human; GLYT-2 transporter; glycine transporter; neurotransmitter;
KW activity: N-methyl-D-aspartate receptor; NMDA;
KW strychnine-sensitive glycine receptor; treatment; dyskinesia;
KW nervous system disorder; myoclonus; muscle spasm; hyperactivity;
KW epilepsy; spasticity; head trauma; neuronal cell death;
KW multiple sclerosis; spinal cord injury; Huntington's disease;
KW amyotrophic lateral sclerosis; ds.
XX
XX Synthetic.
OS Homo sapiens.
FH Key location/qualifiers
FT mutation 2198 /*tag= a /note= "T2198 replaced with C"
FT
FN
PN MO9807854-A1.
PD 26-FEB-1998.
PF 20-AUG-1997; 97WO-US14637.
PR 20-AUG-1996; 96US-0700013.
XX
PA (ALIX) ALLELIX NEUROSCIENCE INC.
PI Albert VR, Borden LA, Kowalski LR, McKelvey JF;
DR WPI; 1998-169158/15.
XX Human glycine transporter - useful for treating nervous system
PT disorders, e.g. pain, myoclonus, etc.
PS Claim 6; Page -: 90pp; English.

The present sequence represents a human GLYT-2 transporter mutant gene sequence. GLYT-2 is a glycine transporter found predominantly in the brain stem and the spinal cord. GLYT-2 regulates the levels of glycine, a major neurotransmitter, that preferentially influences the activity of N-methyl-D-aspartate (NMDA) receptors and strychnine-sensitive glycine receptors. The GLYT-2 protein can be used to treat or to diagnose a nervous system disorder or condition, e.g. pain, myoclonus, muscle spasm or hyperactivity, epilepsy or spasticity associated with stroke, head trauma, neuronal cell death, multiple sclerosis, spinal cord injury, dystonia, Huntington's disease or amyotrophic lateral sclerosis.

note: the present sequence does not appear in the specification; it was created using information provided.

Sequence 2397 BP; 508 A; 666 C; 629 G; 594 T; 0 other;

Query Match 13.7%; Score 149.4; DB 19; Length 2397;
Best Local Similarity 54.9%; Pred No. 4, 2e-28;
Matches 322; Conservative 0; Mismatches 256; Indels 9; Gaps 1

272 GAAGGATCACAAACAATTCAGATTATTATTCCTCAAAGATGAGAAAATGAGGGTT 331

D	b		1750	GAGACCATATGAGACCTCCATCTCAGAGAGATTTCGCCAAGTACTACACACACAAAGCA	1809
O	y		332	CCCATTACCTTTGGGCGTCTGTGGTTTTGTTTCTCCTGGCTCGTGCTGACTGCAG	391
D	b		1810	GTCGTTACTCTGGGCGTCTGCATTTGTTTCTTCATCAGGGTTTCCATCATCTAC	1869
O	y		392	GCTGGAATTTACTCTGGTTCACTCTGATTAACCACCTTCCTGCTGATGGGCATTTA	451
D	b		1870	GGTGGATTTTACAATGTTTCAAGCTTGTTGGACACCATGCTGCTCCTATATGCCCTGTATC	1929
O	y		452	GCACTTACTGAGCTAGTTGATCATCTGATTTATGAGGGAACAGATTCATTGAG	511
D	b		1930	ATTGCCATTTTGTGAGCTCGGGGATCTCTATGTATGTGCTTGCAAAGATTTGTGAA	1989
O	y		512	GATACAGAATAATGATGATTTGGAGCAAGAGGTGATATTCTGCTATGGTGAGAGCTGC	571
D	b		1990	GATTTAGAGATGATGATTTGATTTCCAGCCTAACATCTTC-----TGGAAAGTCTGC	2040
O	y		572	TGGTTTGTAAATTACGCTTACCTTCCTTTGATTGCAATATTATCTGGTCAATTGGCAAT	631
D	b		2041	TGGGCATTTTGAACCCACACCATTTTAACCTTTATCCCTTGCTCAGCTTTTACCAAGTG	2100
O	y		632	CATGACCTAATTTATGGCGCAATTCATACCTCATCGGGAGTTGCTTTAGGGCTGCT	691
D	b		2101	GAGCCCATGACATGATGCGCTTTACCGCTATCTAACCTGCTCATGTCCTGGATGCTA	2160
O	y		692	ATGATTTTCTTCTGCATTTATTTGGATTCACATTTATGGCTATACAAAATTAATTCAGGCT	751
D	b		2161	ATGCTGCGCTGTTCCGTCATCTGATTCACATTTATGCTGTGATTAATAATGCATCTGSC	2220
O	y		752	AAAGGAACATCTTTCAACGCCCTTATAAGTTGCTGCAGACACGACTTCTAAGTGGGCTCA	811
D	b		2221	CCTGGAAGATTTATGAGAGGCTAAAGTTGGTGTGCTGCCACAGCGGACTGGGCCCCA	2280
O	y		812	TACCTGGACACATCGTGGGAAAGATTAACATGATGTAATTCC	858
D	b		2281	TTCTTAGCTCACACCGCGGGAGCGTTACAAAGACATGATGACCC	2327
<hr/>					
RESULT 14					
AAV22932					
ID	AAV22932	standard; cDNA;	2397	BP.	
XX	XX				
AC	AAV22932;				
DT	29-JUL-1998	(first entry)			
DE		DNA encoding a human GlyT-2 transporter protein.			
xx	xx				
KW	Human; GlyT-2 transporter; glycine transporter; neurotransmitter;				
KW	activity; N-methyl-D-aspartate receptor; NMDA;				
KW	strychnine-sensitive glycine receptor; treatment; diagnosis;				
KW	nervous system disorder; myoclonus; muscle spasm; hyperactivity;				
KW	epilepsy; spasticity; head trauma; neuronal cell death;				
KW	multiple sclerosis; spinal cord injury; Huntington's disease;				
KW	amyotrophic lateral sclerosis; ds.				
XX	XX				
OS	Homo sapiens.				
XX	XX				
FH	Key	Location/Qualifiers			
FT	CDS	1..2394			
FT	FT	/tag=			
FT	FT	/product= Glyt-2 transporter			
XX	XX				
PN	WO9807854-A1.				
XX	XX				
PD	26-FEB-1998.				
XX	XX				
PF	20-AUG-1997;	97WO-US14637.			
XX	XX				
PR	20-AUG-1996;	96US-0700013.			
XX	XX				

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 12:07:32 ; Search time 75 Seconds

(Without alignments)
457.133 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198

Sequence: 1 SEETITTTIDLEPKWKWK.....VPRKKGHEIPTVSGSRKPE 216

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1198	100.0	216	23	ABG60919
2	1176	98.2	414	22	AAW23517
3	1176	98.2	642	21	AAV81943
4	1157	96.6	642	21	AAV81944
5	505.5	42.2	797	19	AAW56381
6	503.5	42.0	797	21	AAW56386
7	503.5	42.0	797	21	AAW56387
8	503.5	42.0	797	21	AAW56388
9	503.5	42.0	797	21	AAW56389

10	503.5	42.0	797	21	AAW56390	Human SC polypept
11	502.5	41.9	280	19	AAW56367	Fragment of the hu
12	502.5	41.9	797	19	AAW56372	Amino acid sequenc
13	502.5	41.9	797	19	AAW56373	Human GLYT-2 trans
14	502.5	41.9	797	19	AAW56374	Human GLYT-2 trans
15	502.5	41.9	797	19	AAW56375	Human GLYT-2 trans
16	502.5	41.9	797	19	AAW56376	Human GLYT-2 trans
17	502.5	41.9	797	19	AAW56377	Human GLYT-2 trans
18	502.5	41.9	797	19	AAW56378	Human GLYT-2 trans
19	502.5	41.9	797	19	AAW56379	Human GLYT-2 trans
20	502.5	41.9	797	19	AAW56380	Human GLYT-2 trans
21	502.5	41.9	797	19	AAW56381	Human GLYT-2 trans
22	502.5	41.9	797	19	AAW56382	Human GLYT-2 trans
23	502.5	41.9	797	19	AAW56383	Human GLYT-2 trans
24	502.5	41.9	797	19	AAW56384	Human GLYT-2 trans
25	502.5	41.9	797	19	AAW56385	Human GLYT-2 trans
26	502.5	41.9	797	19	AAW56386	Human GLYT-2 trans
27	502.5	41.9	797	21	AAW56387	Human GLYT-2 trans
28	502.5	41.9	797	21	AAW56388	Human GLYT-2 trans
29	502.5	41.9	797	21	AAW56389	Human GLYT-2 trans
30	502.5	41.9	797	21	AAW56390	Human GLYT-2 trans
31	501.5	41.9	797	21	AAW56391	Human GLYT-2 trans
32	501.5	41.9	797	21	AAW56392	Human GLYT-2 trans
33	501.5	41.9	797	21	AAW56393	Human GLYT-2 trans
34	501.5	41.9	797	21	AAW56394	Human GLYT-2 trans
35	497.5	41.5	797	21	AAW56395	Human GLYT-2 trans
36	496.5	41.4	797	21	AAW56396	Human GLYT-2 trans
37	491.5	41.0	797	21	AAW56397	Human GLYT-2 trans
38	490.5	40.9	797	21	AAW56398	Human GLYT-2 trans
39	462.5	38.6	194	19	AAW56399	Human GLYT-2 trans
40	447.5	37.4	194	19	AAW56400	Human GLYT-2 trans
41	359	30.0	637	18	AAW11704	Fragment of the hu
42	359	30.0	637	18	AAW11705	Fragment of the hu
43	350.5	29.3	150	21	AAW57224	High affinity Na+-
44	350.5	29.3	150	21	AAW57225	Rat prolins transp
45	350.5	29.3	150	21	AAW57226	Human glycine tran

ALIGNMENTS

RESULT 1
ID ABG60919 standard; Protein: 216 AA.
XX ABG60919;
AC
XX 13-AUG-2002 (first entry)
DT
XX Protein encoded by lung specific gene #9.
DE
XX Lung specific gene; gene therapy; vaccine; lung cancer;
KW cancer staging; cancer monitoring; cancer diagnosis;
KW imaging lung cancer; metastases.
XX Homo sapiens.
XX WO200218576-A2.
XX
XX 07-MAR-2002.
XX
XX 27-AUG-2001; 2001WO-US26684.
XX
XX 28-AUG-2000; 2000US-228378P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Chen S, Macina RA, Sun Y, Recipon H;
XX WPI; 2002-434904/46.
XX
XX New lung specific genes and their encoded proteins, useful in gene
XX therapy or as a vaccine for treating lung cancer, as well as for

PT measuring metastases of lung cancer, or staging, monitoring, diagnosing
 PT or imaging lung cancer
 PS Claim 2; Page 205-206; 206pp; English.
 CC The invention describes a new lung specific gene and its variants. The
 CC lung specific gene proteins and genes are useful in gene therapy or as a
 CC vaccine for treating lung cancer. Lung specific genes are also useful for
 CC staging, monitoring, diagnosing or imaging lung cancer, as well as for
 CC measuring metastases of lung cancer. This is the amino acid sequence of
 CC a protein encoded by a lung specific gene described in the invention.
 CC
 SQ Sequence 216 AA;
 Query Match 100.0%; Score 1198; DB 23; Length 216;
 Best Local Similarity 100.0%; Pred. No. 2e-127;
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SEETITTTTIDLPFRVKKMKRPVITLGGCLVFLGLVGYVQAGIYWHLIDHFCAGMI 60
 DB 1 SEETITTTTIDLPFRVKKMKRPVITLGGCLVFLGLVGYVQAGIYWHLIDHFCAGMI 60
 QY 61 LIAAILELVGIIWYIGNGRFIEDTEMMIGAKRWIFMLWWRACWFTVITPILIAIFISLV 120
 DB 61 LIAAILELVGIIWYIGNGRFIEDTEMMIGAKRWIFMLWWRACWFTVITPILIAIFISLV 120
 QY 121 OFHRPNYGAIPYPMGVNLGCMIVFCIIWIPIMAIKIIQAKNIFORLISCCRPASNW 180
 DB 121 OFHRPNYGAIPYPMGVNLGCMIVFCIIWIPIMAIKIIQAKNIFORLISCCRPASNW 180
 QY 181 GPYLEQHRGERYKDMVYPRKEAGHEIPTVSGSRPE 216
 DB 181 GPYLEQHRGERYKDMVYPRKEAGHEIPTVSGSRPE 216
 RESULT 2
 ID AAM23517 standard; Protein; 414 AA.
 AC AAM23517;
 DT 12-OCT-2001 (first entry)
 DE Human EST encoded protein seq ID NO: 1042.
 XX
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.
 OS Homo sapiens.
 PN WO200154477-A2.
 PD 02-AUG-2001.
 PE 25-JAN-2001; 2001WO-US02687.
 PR 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 PA (HSE-) HYSEQ INC.
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 DR WPI: 2001-476164/51.
 DR N-PSDB: AAH98176.
 PT isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use

XX Claim 20; Page 796-797; 1275pp; English.
 PS The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.
 CC
 SQ Sequence 414 AA;
 Query Match 98.2%; Score 1176; DB 22; Length 414;
 Best Local Similarity 98.6%; Pred. No. 1.4e-124;
 Matches 213; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SEETITTTTIDLPFRVKKMKRPVITLGGCLVFLGLVGYVQAGIYWHLIDHFCAGMI 60
 DB 199 SEETITTTTIDLPFRVKKMKRPVITLGGCLVFLGLVGYVQAGIYWHLIDHFCAGMI 258
 QY 61 LIAAILELVGIIWYIGNGRFIEDTEMMIGAKRWIFMLWWRACWFTVITPILIAIFISLV 120
 DB 259 LIAAILELVGIIWYIGNGRFIEDTEMMIGAKRWIFMLWWRACWFTVITPILIAIFISLV 318
 QY 121 OFHRPNYGAIPYPMGVNLGCMIVFCIIWIPIMAIKIIQAKNIFORLISCCRPASNW 180
 DB 319 OFHRPNYGAIPYPMGVNLGCMIVFCIIWIPIMAIKIIQAKNIFORLISCCRPASNW 378
 QY 181 GPYLEQHRGERYKDMVYPRKEAGHEIPTVSGSRPE 216
 DB 379 GPYLEQHRGERYKDMVYPRKEAGHEIPTVSGSRPE 414
 RESULT 3
 ID AAY81943 standard; Protein; 642 AA.
 AC AAY81943;
 DT 30-JUN-2000 (first entry)
 DE Human GUYLIKE protein sequence #1.
 XX
 KW GUYLIKE protein; Glycine transporter protein; human; neuropathic pain;
 KW chronic pain; postoperative pain; rheumatoid arthritis; neuralgia;
 KW neuropathy; algosia; nerve injury; ischaemia; neurodegeneration; stroke;
 KW incontinence; inflammatory disorder; spasticity; myoclonus; epilepsy;
 KW head trauma; chromosome localisation study; therapy.
 OS Homo sapiens.
 PN WO2000014221-A1.
 PD 16-MAR-2000.
 PE 03-SEP-1999; 99WO-GB02909.
 PR 04-SEP-1998; 98GB-0019405.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 PI Duckworth DM, Evans JR;
 DR WPI: 2000-256973/22.
 DR N-PSDB: AAA07378.
 PT New human glycine transporter, known as GUYLIKE, useful in the
 PT treatment of diseases such as neuropathic pain, chronic pain,
 PT neuralgia, ischaemia, stroke, incontinence, spasticity and epilepsy,
 PT comprises a 642 amino acid sequence

PS Claim 2; Page 17-18; 33pp; English.

CC This sequence represents a human glycine transporter (GLYT) protein,
 CC designated GLYTlike. The GLYTlike polypeptide and the polynucleotide
 CC encoding it are useful in the treatment of diseases such as neuropathic
 CC pain, pain, chronic pain, postoperative pain, rheumatoid arthritic pain,
 CC neuralgia, neuropathies, algesia, nerve injury, ischaemia,
 CC neurodegeneration, stroke, incontinence, inflammatory disorders,
 CC spasticity, myoclonus, epilepsy and head trauma. The DNA sequence is
 CC valuable for chromosome localisation studies. The GLYTlike polypeptide
 CC and polynucleotide are also useful in diagnostic assays for detecting
 CC diseases associated with in appropriate GLYTlike activity or levels.

XX Sequence 642 AA;

Query Match 98.2%; Score 1176; DB 21; Length 642;

Best Local Similarity 98.6%; Pred. No. 2.4e-124; Indels 0; Gaps 0;

Matches 213; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SEETITTTTODLPFKVKKKRVPTTGGCLVFLGLGVCYTOAGIYVHLDHFCAGMGI 60
 DB 427 SIETITTTTODLPFKVKKKRVPTTGGCLVFLGLGVCYTOAGIYVHLDHFCAGMGI 486
 QY 61 LIAAILELVGIIWYGGNRFEDTEEMIGAKRMIFWLMWRACWFVTPILLIAIFWSLV 120
 DB 487 LIAAILELVGIIWYGGNRFEDTEEMIGAKRMIFWLMWRACWFVTPILLIAIFWSLV 546
 QY 121 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPMAIIKIIOAKGNIFORLISCCRPASNW 180
 DB 547 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPMAIIKIIOAKGNIFORLISCCRPASNW 606
 QY 181 GPYLEQHRGGRYKDMVVPKKEAGHEIPTVSGSRKPE 216
 DB 607 GPYLEQHRGGRYKDMVVPKKEAGHEIPTVSGSRKPE 642

RESULT 4

AAV81944

ID AAV81944 standard; Protein: 642 AA.

AC AAV81944;

DT 30-JUN-2000 (first entry)

DE Human GLYTlike protein sequence #2.

XX GLYTlike protein; glycine transporter protein; human; neuropathic pain;

KW chronic pain; postoperative pain; rheumatoid arthritic pain; neuralgia;

KW neuropathy; algesia; nerve injury; ischaemia; neurodegeneration; stroke;

KW incontinence; inflammatory disorder; spasticity; myoclonus; epilepsy;

KW head trauma; chromosome localisation study; therapy.

OS Homo sapiens.

XX MO200014221-A1..

PD 16-MAR-2000.

PF 03-SEP-1999; 99MO-GB02909.

PR 04-SEP-1998; 98GB-0019405.

PA (SMIK) SMITHKLINE BEECHAM PLC.

XX Duckworth DM, Evans JR;

DR WPI; 2000-256973/22.

DR N-PSDB; AAA07379.

XX New human glycine transporter, known as GLYTlike, useful in the

PT treatment of diseases such as neuropathic pain, chronic pain,

PT neuralgia, ischaemia, stroke, incontinence, spasticity and epilepsy,

PT comprises a 642 amino acid sequence -

XX Claim 12; Page 19; 33pp; English.

CC This sequence represents a human glycine transporter (GLYT) protein,
 CC designated GLYTlike. The GLYTlike polypeptide and the polynucleotide
 CC encoding it are useful in the treatment of diseases such as neuropathic
 CC pain, pain, chronic pain, postoperative pain, rheumatoid arthritic pain,
 CC neuralgia, neuropathies, algesia, nerve injury, ischaemia,
 CC neurodegeneration, stroke, incontinence, inflammatory disorders,
 CC spasticity, myoclonus, epilepsy and head trauma. The DNA sequence is
 CC valuable for chromosome localisation studies. The GLYTlike polypeptide
 CC and polynucleotide are also useful in diagnostic assays for detecting
 CC diseases associated with in appropriate GLYTlike activity or levels.

XX Sequence 642 AA;

Query Match 96.6%; Score 1157; DB 21; Length 642;

Best Local Similarity 96.3%; Pred. No. 3.4e-122; Indels 0; Gaps 0;

Matches 208; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SEETITTTTODLPFKVKKKRVPTTGGCLVFLGLGVCYTOAGIYVHLDHFCAGMGI 60
 DB 427 SIETITTTTODLPFKVKKKRVPTTGGCLVFLGLGVCYTOAGIYVHLDHFCAGMGI 486
 QY 61 LIAAILELVGIIWYGGNRFEDTEEMIGAKRMIFWLMWRACWFVTPILLIAIFWSLV 120
 DB 487 LIAAILELVGIIWYGGNRFEDTEEMIGAKRMIFWLMWRACWFVTPILLIAIFWSLV 546
 QY 121 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPMAIIKIIOAKGNIFORLISCCRPASNW 180
 DB 547 QFHRPNYGAIPYDPMGVALGCMIVFCIIWIPMAIIKIIOAKGNIFORLISCCRPASNW 606
 QY 181 GPYLEQHRGGRYKDMVVPKKEAGHEIPTVSGSRKPE 216
 DB 607 GPYLEQHRGGRYKDMVVPKKEAGHEIPTVSGSRKPE 642

RESULT 5

AAW56381

ID AAW56381 standard; Protein: 797 AA.

AC AAW56381;

DT 29-JUL-1998 (first entry)

DE Human GLYT-2 transporter protein F733S mutant.

XX Human; GLYT-2 transporter; glycine transporter; neurotransmitter;

KW activity; N-methyl-D-aspartate receptor; NMDA;

KW strychnine-sensitive glycine receptor; treatment; diagnosis;

KW nervous system disorder; myoclonus; muscle spasm; hyperactivity;

KW multiple sclerosis; head trauma; neuronal cell death;

KW amyotrophic lateral sclerosis.

OS Synthetic.

XX Homo sapiens.

OS Homo sapiens.

FT Key location/Qualifiers

FT MISC-difference 733 /label= F733S

FT /note= "Phe replaced with Ser"

XX WO9807854-A1.

XX 26-FEB-1998.

XX 20-AUG-1997; 97WO-US14637.

XX 20-AUG-1996; 96US-0700013.

XX (ALLX) ALLELIX NEUROSCIENCE INC.

XX WPI: 2000-387780/33.
 DR N-PSDB: AAA61440.
 XX
 PT New nucleic acid encoding a human glycine transporter type 2 (GLYT2),
 PT useful for expressing GLYT2 which can then be used for detecting
 PT whether an animal has autoimmune antibodies against a glycine
 PT transporter -
 XX
 PS Claim 10; Page 167-169; 175pp; English.
 XX
 CC The present sequence is the consensus protein sequence of the human
 CC glycine transporter type 2 (GLYT2). The researchers sequenced a number of
 CC clones of the sequence from spinal cord and astrocytoma cells to produce
 CC this consensus version. Glycine is involved in neurotransmission in the
 CC spinal cord and cerebellum, where glycine receptor inhibition results in
 CC pain transmission. The GLYT2 gene and protein could, therefore, be used
 CC to identify inhibitory compounds which prevent GLYT2 activity, causing a
 CC decrease in muscle hyperactivity. They could be used to treat diseases or
 CC conditions associated with muscle contraction, including muscle
 CC spasticity (such as that due to epilepsy, stroke, head trauma, multiple
 CC sclerosis, neuronal cell death, spinal cord injury and dystonia),
 CC myoclonus, epilepsy, Huntington's disease, amyotrophic lateral sclerosis
 CC and pain.
 CC
 SO Sequence 797 AA:
 Query Match 42.0%; Score 503.5; DB 21; Length 797;
 Best Local Similarity 46.5%; Pred. No. 5e-48;
 Matches 92; Conservative 33; Mismatches 70; Indels 3; Gaps 1;
 QY 1 SEETITTTIDLEPKVKKKRRVPTTGCCVLFLGLVCTQAGIYVWHLIDHFCAGWI 60
 Db 582 SIETIYTSIDEPKRYRTKRPVTLGCCCFIFMGPMITGGIYFQVLDYTAASYAL 641
 QY 61 LIAAIELVGIWVYGNRFIEDTEMIGAKRMIFWLMRACWFTVITPILIAIFWISLV 120
 Db 642 VVIAIFELVGISYVYGLORCEDIEIMIGOPNIF--WKVCNAFVPTITLFLCSFY 698
 QY 121 QFHRPNYGAIPYDNGVALGCMIVFCIIPIMAIKITIOAGNIFORLISCRPASNW 180
 Db 699 QMEPMYGSYRYPNMSVGLMLACSVIPIPIMFVYIKMLADGRFLERLKLVCSPQPDW 758
 QY 181 GPYLEQHRGERYKDMVVP 198
 Db 759 GPFLAQHRGERYKKNMIDP 776
 Db
 RESULT 8
 AAB23326 standard; Protein: 797 AA.
 XX
 AC AAB23326;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Human glycine transporter type 2 protein mutant #8.
 XX
 KW Human glycine transporter type 2; GLYT2; muscle spasticity; myoclonus;
 KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;
 KW dystonia; Huntington's disease; amyotrophic lateral sclerosis;
 KW mutant; mutlein.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 582 /note="wild-type Thr substituted by Ser"
 XX
 PN WO200029564-A2.
 XX
 PD 25-MAY-2000.

XX 09-NOV-1999; 99WO-US26534.
 PF
 XX 12-NOV-1998; 98US-0191468.
 PR
 XX (GLIA-) GLINTECH INC.
 PA
 PI Gallagher MJ, Burgess LH, Brunden KR;
 XX WPI: 2000-387780/33.
 DR
 XX
 PT New nucleic acid encoding a human glycine transporter type 2 (GLYT2),
 PT useful for expressing GLYT2 which can then be used for detecting
 PT whether an animal has autoimmune antibodies against a glycine
 PT transporter -
 XX
 PS Claim 7; Page -; 175pp; English.
 XX
 CC The present sequence is a mutant version of the protein sequence of the
 CC human glycine transporter type 2 (GLYT2). Glycine is involved in
 CC neurotransmission in the spinal cord and cerebellum, where glycine
 CC receptor inhibition results in pain transmission. The GLYT2 gene and
 CC protein could, therefore, be used to identify inhibitory compounds which
 CC prevent GLYT2 activity, causing a decrease in muscle hyperactivity. They
 CC could be used to treat diseases or conditions associated with muscle
 CC contraction, including muscle spasticity (such as that due to epilepsy,
 CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal
 CC cord injury and dystonia), myoclonus, epilepsy, Huntington's disease,
 CC amyotrophic lateral sclerosis and pain.
 CC Note: This sequence is not shown in the specification, but is derived
 CC from the hGLYT2 wild type sequence shown in SEQ ID NO: 124 (AAB09898).
 CC
 SO Sequence 797 AA:
 Query Match 42.0%; Score 503.5; DB 21; Length 797;
 Best Local Similarity 46.5%; Pred. No. 5e-48;
 Matches 92; Conservative 33; Mismatches 70; Indels 3; Gaps 1;
 QY 1 SEETITTTIDLEPKVKKKRRVPTTGCCVLFLGLVCTQAGIYVWHLIDHFCAGWI 60
 Db 582 SIETIYTSIDEPKRYRTKRPVTLGCCCFIFMGPMITGGIYFQVLDYTAASYAL 641
 QY 61 LIAAIELVGIWVYGNRFIEDTEMIGAKRMIFWLMRACWFTVITPILIAIFWISLV 120
 Db 642 VVIAIFELVGISYVYGLORCEDIEIMIGOPNIF--WKVCNAFVPTITLFLCSFY 698
 QY 121 QFHRPNYGAIPYDNGVALGCMIVFCIIPIMAIKITIOAGNIFORLISCRPASNW 180
 Db 699 QMEPMYGSYRYPNMSVGLMLACSVIPIPIMFVYIKMLADGRFLERLKLVCSPQPDW 758
 QY 181 GPYLEQHRGERYKDMVVP 198
 Db 759 GPFLAQHRGERYKKNMIDP 776
 Db
 RESULT 9
 AAB23327 standard; Protein: 797 AA.
 XX
 AC AAB23327;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Human glycine transporter type 2 protein mutant #9.
 XX
 KW Human glycine transporter type 2; GLYT2; muscle spasticity; myoclonus;
 KW epilepsy; stroke; head trauma; multiple sclerosis; spinal cord injury;
 KW dystonia; Huntington's disease; amyotrophic lateral sclerosis;
 KW mutant; mutlein.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX

Key Location/Qualifiers

FT Misc-difference 24 /note- "wild-type Gly substituted by Ser"

FT Misc-difference 74 /note- "wild-type Arg substituted by Trp"

FT Misc-difference 155 /note- "wild-type Ser substituted by Gly"

FT Misc-difference 188 /note- "wild-type Asn substituted by Asp"

FT Misc-difference 362 /note- "wild-type Gln substituted by Leu"

FT Misc-difference 431 /note- "wild-type Val substituted by Ala"

FT Misc-difference 582 /note- "wild-type Thr substituted by Ser"

XX WO200029564-A2.

XX 25-MAY-2000.

XX 09-NOV-1999; 99WO-US26534.

XX 12-NOV-1998; 98US-0191468.

XX (GLIA-) GLIATECH INC.

XX Gallagher MJ, Burgess LH, Brunden KR.

XX WPI; 2000-387780/33.

XX New nucleic acid encoding a human glycine transporter type 2 (GLYT2),

PT useful for expressing GLYT2 which can then be used for detecting

PT whether an animal has autoimmune antibodies against a glycine

PT transporter -

XX Claim 8; Page -: 175bp; English.

XX The present sequence is a mutant version of the protein sequence of the

CC human glycine transporter type 2 (GLYT2). Glycine is involved in

CC neurotransmission in the spinal cord and cerebellum where glycine

CC receptor inhibition results in pain transmission. The GLYT2 gene and

CC protein could, therefore, be used to identify inhibitory compounds which

CC prevent GLYT2 activity, causing a decrease in muscle hyperactivity. They

CC could be used to treat diseases or conditions associated with muscle

CC contraction, including muscle spasticity (such as that due to epilepsy,

CC stroke, head trauma, multiple sclerosis, neuronal cell death, spinal

CC cord injury and dystonia), myoclonus, epilepsy, Huntington's disease,

CC amyotrophic lateral sclerosis and pain.

CC Note: This sequence is not shown in the specification, but is derived

CC from the hGLYT2 wild type sequence shown in SEQ ID NO: 124 (AAB09898).

XX

SO Sequence 797 AA;

Query Match 42.0%; Score 503.5; DB 21; Length 797;

Best Local Similarity 46.5%; Pred. No. 5e-48;

Matches 92; Conservative 33; Mismatches 70; Indels 3; Gaps 1;

QY 1 SEETITTTODLPPKMKMRVPITLGGCLVFLGLGYCTOAGIYVHLIDHRCAGMGI 60

DB 582 SIEITVTSISDEPKLTKHKKPVFTLGGCICFFIMGGITMGGIYQVLDVTAASVAL 641

QY 61 LIAAILELGIIVYGGNFIETDEMIGAKRWIFWLMWRACWFTPTILLIAIFWISLV 120

DB 642 VITAIPELVGISVYVIGRCEDEIMKIGOPNIF---WKVCMAFVPTITLFTILCFSEY 698

QY 121 QEHHPNNGAIPYDVGVALGCMIVFCIWIPIMAIKITIOAKGNIFORLISCCRPASNM 180

DB 699 QMEPMYGYGYRYPNMSVGLMLACSVIWIPIFVIMKHLAPGRFIERLKLCSPOPDW 758

QY 181 GPYLEOHRGERYKDMVVP 198

DB 759 GPFLAHRGERYKDMVVP 776

RESULT 10

ID AAY82115 standard; protein; 797 AA.

XX AAY82115;

XX 05-JUN-2000 (first entry)

DE Human SC polypeptide glycine transporter SEQ ID NO:2.

XX Human; glycine transporter; SC6 polypeptide; screening; diagnosis;

KW identification; neuropathic pain; spasticity; myoclonus; epilepsy;

KW stroke; head trauma; spinal cord injury; dystonia; multiple sclerosis;

KW amyotrophic lateral sclerosis; Huntington's disease;

KW Parkinson's disease.

XX Homo sapiens.

XX EP982400-A1.

XX 01-MAR-2000.

XX 27-OCT-1998; 98EP-0203615.

XX 28-AUG-1998; 98GB-0018890.

XX (SMK) SMITHKLINE BEECHAM PLC.

XX Brown AM, Chapman CG, Gloger IS, Evans JR, Cairns W, Herton H;

XX WPI; 2000-258613/23.

XX N-PSDB; AA295536.

XX New isolated polypeptides and polynucleotides, useful for screening for

PT (ant)agonists and for diagnosing a disease e.g. spasticity, epilepsy,

PT stroke, multiple sclerosis, Huntington's disease or Parkinson's disease

XX Claim 4; Page 17-20; 27bp; English.

XX The present sequence represents the human SC polypeptide glycine

CC transporter. The polypeptide is useful in a method for screening to

CC identify compounds which stimulate or inhibit the function of the

CC polypeptide and in a process for diagnosing a disease or a

CC susceptibility to a disease in a subject related to expression or

CC activity of the polypeptide e.g. treatment of neuropathic pain, other

CC pain, spasticity, myoclonus, epilepsy, stroke, head trauma, spinal cord

CC injury, dystonia, multiple sclerosis, amyotrophic lateral sclerosis,

CC Huntington's disease or Parkinson's disease.

XX

SO Sequence 797 AA;

Query Match 42.0%; Score 503.5; DB 21; Length 797;

Best Local Similarity 46.4%; Pred. No. 5e-48;

Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1;

QY 3 ETTTITTTODLPPKMKMRVPITLGGCLVFLGLGYCTOAGIYVHLIDHRCAGMGI 62

DB 584 ETTVTSISDEPKLTKHKKPVFTLGGCICFFIMGGITMGGIYQVLDVTAASVALYI 643

QY 63 AAILLELGIIVYGGNFIETDEMIGAKRWIFWLMWRACWFTPTILLIAIFWISLV 122

DB 644 IAIPELVGISVYVIGRCEDEIMKIGOPNIF---WKVCMAFVPTITLFTILCFSEY 700

QY 123 HRENYGAIPYDVGVALGCMIVFCIWIPIMAIKITIOAKGNIFORLISCCRPASNM 182

DB 701 EPMYGYGYRYPNMSVGLMLACSVIWIPIFVIMKHLAPGRFIERLKLCSPOPDW 760

QY 183 YLEOHRGERYKDMVVP 198

DB 761 FLAHRGERYKDMVVP 776

[illegible]

DB 644 IAI FELV G I S V Y G L O R F C E D I E M M I G F O P N I F - - - W K V C W A F V T P T I L T F I L C F S F Y Q W 70

DB 644 IAI FELV G I S V Y G L O R F C E D I E M M I G F O P N I F - - - W K V C W A F V T P T I L T F I L C F S F Y Q W 70

OY 123 HRPNGAIDPDMGVALGCMIVFCITWIPIMAIITIIIOAKGNIFORLISCCRPASNMGP 182
 DB 701 EPMYGSYRYPNWSMVLGMLACSVIPIWIMFVIMKMLAPGRFTRKLVCSPQDWGP 760
 OY 183 YLEOHNGERYKDMVP 198
 DB 761 FLAÖHNGERYKMDP 776

RESULT 13

AAM56373
 ID AAM56373 standard; protein; 797 AA.

AC AAM56373;
 DT 29-JUL-1998 (first entry)

DE Human GLYT-2 transporter protein G102S mutant.

KW Human; GLYT-2 transporter; glycine transporter; neurotransmitter;
 KW activity; N-methyl-D-aspartate receptor; NMDA;
 KW strychnine-sensitive glycine receptor; treatment; diagnosis;
 KW nervous system disorder; myoclonus; muscle spasm; hyperactivity;
 KW epilepsy; spasticity; head trauma; neuronal cell death;
 KW multiple sclerosis; spinal cord injury; Huntington's disease;
 KW amyotrophic lateral sclerosis.

OS Synthetic.
 OS Homo sapiens.

FT Key Location/Qualifiers
 FT Misc-difference 102 /label=G102S
 FT /note="Gly replaced with Ser"

XX WO9807854-A1.

XX 26-FEB-1998.

XX 20-AUG-1997; 97WO-US14637.

XX 20-AUG-1996; 96US-0700013.

XX (ALIX) ALLELIX NEUROSCIENCE INC.

XX PI Albert VR, Borden LA, Kowalski LR, McKelvy JF;

XX WPI; 1998-169158/15.

XX Human glycine transporter - useful for treating nervous system
 PT disorders, e.g. pain, myoclonus, etc.

XX Claim 1; Page -: 90pp; English.

XX The present sequence represents a human GLYT-2 transporter mutant
 CC protein. GLYT-2 is a glycine transporter found predominantly in
 CC the brain stem and the spinal cord. GLYT-2 regulates the levels of
 CC glycine, a major neurotransmitter, that preferentially influences the
 CC activity of N-methyl-D-aspartate (NMDA) receptors and
 CC strychnine-sensitive glycine receptors. The GLYT-2 protein can be
 CC used to treat or to diagnose a nervous system disorder or condition,
 CC e.g. pain, myoclonus, muscle spasm or hyperactivity, epilepsy or
 CC spasticity associated with stroke, head trauma, neuronal cell death,
 CC multiple sclerosis, spinal cord injury, dystonia, Huntington's disease
 CC or amyotrophic lateral sclerosis.
 CC note: the present sequence does not appear in the specification; it was
 CC created using information provided.

XX Sequence 797 AA;

Query Match 41.9%; Score 502.5; DB 19; Length 797;
 Best Local Similarity 46.4%; Pred. No. 6.5e-46;

Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1;
 OY 3 EMTITTTIDLPFRKAKKAVPTITGCCVLVFLGLVCYQAGIYVWHILIDHFCAGMGLI 62
 DB 584 ETIVTSIDSEFPKRYLTKHPVPTLCCICCEFTMGPMPTGGGIMFQLDVITAAAYLVI 643
 OY 63 AALIELVGIIWITGNGRFTEDTMMIGARWTFWIMWRACVPTIPILLIAFIWSLVQF 122
 DB 644 IAFELVGISYVYGIDRFCEDEIMMIGFQPNIF---WKVCMAFVPTITFTLCFSFYQW 700
 OY 123 HRPNGAIDPDMGVALGCMIVFCITWIPIMAIITIIIOAKGNIFORLISCCRPASNMGP 182
 DB 701 EPMYGSYRYPNWSMVLGMLACSVIPIWIMFVIMKMLAPGRFTRKLVCSPQDWGP 760
 OY 183 YLEOHNGERYKDMVP 198
 DB 761 FLAÖHNGERYKMDP 776

RESULT 14

AAM56374
 ID AAM56374 standard; protein; 797 AA.

AC AAM56374;
 DT 29-JUL-1998 (first entry)

DE Human GLYT-2 transporter protein S124G mutant.

KW Human; GLYT-2 transporter; glycine transporter; neurotransmitter;
 KW activity; N-methyl-D-aspartate receptor; NMDA;
 KW strychnine-sensitive glycine receptor; treatment; diagnosis;
 KW nervous system disorder; myoclonus; muscle spasm; hyperactivity;
 KW epilepsy; spasticity; head trauma; neuronal cell death;
 KW multiple sclerosis; spinal cord injury; Huntington's disease;
 KW amyotrophic lateral sclerosis.

OS Synthetic.
 OS Homo sapiens.

FT Key Location/Qualifiers
 FT Misc-difference 124 /label=S124G
 FT /note="Ser replaced with Gly"

XX WO9807854-A1.

XX 26-FEB-1998.

XX 20-AUG-1997; 97WO-US14637.

XX 20-AUG-1996; 96US-0700013.

XX (ALIX) ALLELIX NEUROSCIENCE INC.

XX PI Albert VR, Borden LA, Kowalski LR, McKelvy JF;

XX WPI; 1998-169158/15.

XX Human glycine transporter - useful for treating nervous system
 PT disorders, e.g. pain, myoclonus, etc.

XX Claim 1; Page -: 90pp; English.

XX The present sequence represents a human GLYT-2 transporter mutant
 CC protein. GLYT-2 is a glycine transporter found predominantly in
 CC the brain stem and the spinal cord. GLYT-2 regulates the levels of
 CC glycine, a major neurotransmitter, that preferentially influences the
 CC activity of N-methyl-D-aspartate (NMDA) receptors and
 CC strychnine-sensitive glycine receptors. The GLYT-2 protein can be
 CC used to treat or to diagnose a nervous system disorder or condition,
 CC e.g. pain, myoclonus, muscle spasm or hyperactivity, epilepsy or
 CC spasticity associated with stroke, head trauma, neuronal cell death,

Sequence 797 AA;

conservative 33; Mismatches 69; Indels 3; Gaps 1;

3 **ELIIVIT**IO~~DL~~PRVVMKKRRVPI~~LG~~CCVLV~~FL~~LGVCV~~TO~~AGIYV~~W~~HLIDHFCAGWGILI 62

63 AALLEVGIMVGNRFIEDTEMIGAKRWIFWLMWDCMEVYTPILLATFTSTVDE 123
QYVGGILFQLVDVYASVAVI 643

644 IAFELVGISVYGLRFCEDEIEMIGQPNI---MKVCMAEVTPTILTFILCFSEFYQM 700

123 **INPRTGALPIYDGMGVALGMCMLVECLIMIPIMAIILIKITQAKGNIFORTLISCCRPASNNGP** 182

183 YLEOHGGERYKDMVP 198
..... ONE LEMNLVCSFQPDGMP /60

761 FLAQRGGRYKNMIDP 776

RESULT 15
AAW56375

AAW56375: standard; Protein; 797 AA.

29-JUL-1998 (first entry)

Human GLYT-2 transporter protein I279N mutant.

glycine transporter; neurotransmitter; activity; N-methyl-D-aspartate receptor; strychnine-sensitive glycine receptor; NMDA;

nervous system disorder; myoclonus; muscle spasm; hyperactivity; epilepsy; spasticity; head trauma; neuronal cell death; multiple sclerosis

amyotrophic lateral sclerosis.

Homo sapiens.

Location/Qualifiers	279
Misc-difference	

```
-----
T      /note= "file replaced with Asn"
X      -----
XXXXXXXXXXXXXXXXXXXX
```

26-FEB-1998.

20-AUG-1997; 97MO-US14637.

20-AUG-1996; 96US-0700013.

ALBERT VR, BORDEN LA, KOWALSKI JR. MCKELWY TR.

WPI, 1998-169158/15.

Human glycine transporter - useful for treating nervous system disorders, e.g. pain, myoclonus, etc.

Claim 1; Page -; 90pp; English.

Sequence 797 AA;

Query Match	41.9%	Score 502.5;	DB 19;	Length 797;
Best Local Similarity	46.4%	Pred. No. 6	5e+48.	

	Indels	Gaps
Mismatches	69,	1
Conservative	33;	
ETTTTTTODLPPVKKKKRBVDPTMCCCTVRRCAAGCAGACAAAATTTAAAAA OY		

Db

584 ETIVTISDFPKYLRTHKPVFTLLGCCICFIFNGFPMPTGGIIMFOLDVTAASYAIVT 643

63 AALLEVGIIWYGGNRFIEDTEMMIGAKRWFIMLRACWEVITPILLIAIFWISLVQF 122

123 HRPNNGAIPYDDMGVALGWCMTVECTETIRDTATV-----KRCVMAFVPTLTITFLCFSEYQW 700

Ddb
701 EPMTGSGYRYPNMSWMLGLMLACSYIWPPIFVVKHHLAPGFIERKLVCSPODDMGP 760

183 YLEQHRGERYKDMVVP 198

[illegible]

Search completed: August 27, 2003, 17:34:03
Job time : 77 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 17:27:24 ; Search time 37 Seconds

(without alignments)
561.417 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198
Sequence: 1 SEETTTTITQDLFPKVRKM.....VPKKEAGHEIPTVSGSRKPE 216

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	504.5	42.1	799	2 A48716	glycine transport
2	359	30.0	637	2 JH0674	L-proline transport
3	345.5	28.8	638	2 I77912	glycine transport
4	345.5	28.8	692	2 I57956	glycine transport
5	344.5	28.8	633	2 I58140	glycine transport
6	344.5	28.8	638	2 JH0673	glycine transport
7	339.5	28.3	633	2 S23151	glycine transport
8	327	27.3	619	2 S20346	dopamine transport
9	327	27.3	619	2 I59358	dopamine transport
10	319	26.6	620	2 A48960	dopamine transport
11	319	26.6	620	2 I57937	dopamine transport
12	315	26.3	693	2 A41617	dopamine transport
13	308.5	25.8	598	2 F46027	dopamine transport
14	308.5	25.8	598	1 ACRNCT	gamma-aminobutyric
15	300.5	25.1	622	2 S63559	GABA/beta-alanine
16	299	25.0	617	2 S14278	noradrenaline tran
17	296	24.7	599	2 S11073	gamma-aminobutyric
18	288	24.0	598	2 I51368	gamma-aminobutyric
19	286.5	23.9	627	2 JH0695	gamma-aminobutyric
20	286	23.9	615	2 S43285	noradrenaline tran
21	283.5	23.7	602	2 A44409	gamma-aminobutyric
22	283.5	23.7	602	2 A45078	gamma-aminobutyric
23	281.5	23.5	559	2 F88392	protein M01G5.5 [1
24	281.5	23.5	559	2 T33425	hypothetical prote
25	278.5	23.2	627	2 B44409	gamma-aminobutyric
26	273	22.8	614	2 A41757	betaine transport
27	269	22.5	602	2 I55651	noradrenaline tran
28	265.5	22.2	611	2 S46260	creatine transport
29	263	22.0	614	2 S68236	betaine/GABA trans

30	260.5	21.7	620	2 S46487	taurine transport
31	260.5	21.7	620	2 G01426	taurine transport
32	258.5	21.6	655	2 A46270	Na and Cl dependen
33	256.5	21.4	635	2 A46061	Na(+)-dependent cr
34	255.5	21.3	621	2 I57939	taurine transport
35	252.5	21.1	614	2 A43390	gamma-aminobutyric
36	252.5	21.1	635	2 J02386	creatine transport
37	252.5	21.1	635	2 G02095	choleline transport
38	251.5	21.0	635	2 S23431	serotonin transport
39	249.5	20.8	607	2 S19585	neurotransmitter t
40	249.5	20.8	630	2 S30604	GABA transport pro
41	246	20.5	597	2 S65673	taurine transport
42	244	20.4	619	2 S29839	serotonin transport
43	243.5	20.3	630	2 A47398	creatine transport
44	239.5	20.0	426	2 G02277	hypothetical prote
45	234	19.5	632	2 T24405	

ALIGNMENTS

RESULT 1

A48716

glycine transporter GLYT2 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 02-Jun-1995

C:Accession: A48716

R:Lin, Q.R.: Lopez-Corcuera, B.; Mandlyan, S.; Nelson, H.; Nelson, N.

J. Biol. Chem. 268, 22802-22808, 1993

A>Title: Cloning and expression of a spinal cord- and brain-specific glycine transport

A:Reference number: A48716; MUID:94043045; PMID:8226790

A:Accession: A48716

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-799 <LIU>

A:Cross-references: GB:L21672

C:Genetics:

A:Gene: GLYT2

C:Keywords: brain; transmembrane protein

Query Match

Best Local Similarity 42.1%; Score 504.5; DB 2; Length 799;

Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1;

QY 3 ETTTITQDLFPKVRKMRRPITLGGCIVLFLGLVCVQAGIYVWHLIDHFCAGGILI 62

Db 586 ETTTISIDFPKYLRTKRVFTLGGCICFFIFPMITOGGYMFQLDVYAAASYALVI 645

QY 63 AALIELVGIITWYGNRFEDTEMTGAKRMIFWLMWRACWFYIPITLITAFIMSLVGF 122

Db 646 IALFELVGISYVGLQRFEDIEIMIGFQPNIF---WKVCMAFVPTLITFLICRSFYQW 702

QY 123 HRNNGAIPYPMGVALLGCMYFCILIMPIAIIITIOAKNIFQRLISCRPASNMGP 182

Db 703 EPTTYSYRPMNSWVGLMLACSVIWMIFVIMVILAPRFTERTLVCSPPQDMGP 762

QY 183 YLEOHGRERYKDMVVP 198

Db 763 FLAQRGRERYKNMIDP 778

RESULT 2

JH0674

L-proline transport protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 28-Feb-1997

C:Accession: JH0674

R:Freneau Jr., R.T.; Caron, M.G.; Blakely, R.D.

Neuron 8, 915-926, 1992

A>Title: Molecular cloning and expression of a high affinity L-proline transporter ex

A:Reference number: JH0674; MUID:92265305; PMID:1350201

A:Accession: JH0674

A:Molecule type: mRNA

Qy 61 LIAALLELVGIIMYIGNRFIEDTEMMIGAKRMIFMLWRACMFVITPILLIAIFWISLV 120
 Db 464 VVICSCIMCVSIMYIGHRNFYQDIOIMLGFPPPLF---FQICMRFSPTIIFLLIFTVI 520
 Qy 121 QFHRPNYGAIPYDWCVALGCMIVFCIIMPIIMAIKIITQAKN-IFQRLISCCRPASN 179
 Db 521 QYRPITYNHQYCGMAVAIGFLMASSVICIPYALFQLCRTDGLTLQRLKNAKTRPSRD 580
 Qy 180 WGPYLEQHRGERYKDMVVPKKEAGHEI 206
 Db 581 WGPALLEHRTGRVAPTTTSPEDGEFV 607

RESULT 6

glycine transporter protein - rat
 JH0673
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 20-Aug-1999
 C:Accession: JH0673; A46204
 R:Smith, K.E.; Borden, L.A.; Hartig, P.R.; Brancheck, T.; Welshank, R.L.
 A:Title: Cloning and expression of a glycine transporter reveal colocalization with NMDA
 A:Reference number: JH0673; MUID:92265306; PMID:1534013
 A:Accession: JH0673
 A:Molecule type: mRNA
 A:Residues: 1-638 <SMI>
 A:Cross-references: GB:M88595; NID:g204433; PIDN:AAA41257.1; PID:g204434
 A:Experimental source: brain
 R:Gustella, J.; Brecha, N.; Weigmann, C.; Lester, H.A.; Davidson, N.
 Proc. Natl. Acad. Sci. U.S.A. 89, 7189-7193, 1992
 A:Title: Cloning, expression, and localization of a rat brain high-affinity glycine tran
 A:Reference number: A46204; MUID:9235797; PMID:1353889
 A:Accession: A46204
 A:Status: preliminary; not compared with conceptual translation
 A:Residues: 'M','3','GK','6','AKGML','16-638 <GUA>
 A:Molecule type: nucleic acid
 A:Cross-references: GB:M95413; NID:g204431; PIDN:AAA41256.1; PID:g204432
 A:Experimental source: C6 glioma cells and tissues
 A:Note: Sequence extracted from NCBI backbone (NCBI:P:113358)
 C:Superfamily: gamma-aminobutyric acid transporter
 C:Keywords: glycoprotein; transmembrane protein
 F:41-61/Domain: transmembrane #status predicted <TM1>
 F:68-88/Domain: transmembrane #status predicted <TM1>
 F:112-132/Domain: transmembrane #status predicted <TM3>
 F:222-242/Domain: transmembrane #status predicted <TM4>
 F:247-267/Domain: transmembrane #status predicted <TM5>
 F:300-320/Domain: transmembrane #status predicted <TM6>
 F:329-349/Domain: transmembrane #status predicted <TM7>
 F:382-402/Domain: transmembrane #status predicted <TM8>
 F:430-450/Domain: transmembrane #status predicted <TM9>
 F:466-486/Domain: transmembrane #status predicted <TM10>
 F:502-522/Domain: transmembrane #status predicted <TM11>
 F:544-564/Domain: transmembrane #status predicted <TM12>
 F:169-172,182,188/Blinding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 28.8%; Score 344.5; DB 2; Length 638;
 Matches 67; Conservative 54; Mismatches 79; Indels 7; Gaps 4;

Qy 3 ETTITTTIODELPK--VVKMKRVPTLGGCLVFLGLVCTQAGIYVWHLIDHFCAGWGI 60
 Db 410 ETLVTAIVDEVGEMWIK-KTYVTLGVAVAGFLGLPTLSQAGIYVWHLMDYNAVSFSL 468
 Qy 61 LIAALLELVGIIMYIGNRFIEDTEMMIGAKRMIFMLWRACMFVITPILLIAIFWISLV 120
 Db 469 VVICSCIMCVSIMYIGHRNFYQDIOIMLGFPPPLF---FQICMRFSPTIIFLLIFTVI 520
 Qy 121 QFHRPNYGAIPYDWCVALGCMIVFCIIMPIIMAIKIITQAKN-IFQRLISCCRPASN 179
 Db 526 QYRPITYNHQYCGMAVAIGFLMASSVICIPYALFQLCRTDGLTLQRLKNAKTRPSRD 585
 Qy 180 WGPYLEQHRGERYKDMVVPKKEAGHEI 206
 Db 581 WGPALLEHRTGRVAPTTTSPEDGEFV 607

Db 586 WGPALLEHRTGRVAPTTTSPEDGEFV 612

RESULT 7

glycine transporter protein GLYT1a - mouse
 S23151
 C:Species: Mus musculus (house mouse)
 C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
 C:Accession: S23151; B48716
 R:Li, Q.R.; Nelson, H.; Mandliyan, S.; Lopez-Corcuera, B.; Nelson, N.
 FEBS Lett. 305, 110-114, 1992
 A:Title: Cloning and expression of a glycine transporter from mouse brain.
 A:Reference number: S23151; MUID:92316254; PMID:1618338
 A:Accession: S23151
 A:Molecule type: mRNA
 A:Residues: 1-633 <LIU>
 A:Cross-references: EMBL:X67056; NID:951092; PIDN:CAA47440.1; PID:951093
 R:Li, Q.R.; Lopez-Corcuera, B.; Mandliyan, S.; Nelson, H.; Nelson, N.
 J. Biol. Chem. 268, 22802-22808, 1993
 A:Title: Cloning and expression of a spinal cord- and brain-specific glycine transpor
 A:Reference number: A48716; MUID:94043045; PMID:8226790
 A:Accession: B48716
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-62 <LI2>
 A:Genetics:
 C:Superfamily: gamma-aminobutyric acid transporter
 C:Keywords: alternative splicing; transmembrane protein

Query Match

Best Local Similarity 28.3%; Score 339.5; DB 2; Length 633;
 Matches 66; Conservative 54; Mismatches 80; Indels 7; Gaps 4;

Qy 3 ETTITTTIODELPK--VVKMKRVPTLGGCLVFLGLVCTQAGIYVWHLIDHFCAGWGI 60
 Db 405 ETLVTAIVDEVGEMWIK-KTYVTLGVAVAGFLGLPTLSQAGIYVWHLMDYNAVSFSL 463
 Qy 61 LIAALLELVGIIMYIGNRFIEDTEMMIGAKRMIFMLWRACMFVITPILLIAIFWISLV 120
 Db 464 VVICSCIMCVSIMYIGHRNFYQDIOIMLGFPPPLF---FQICMRFSPTIIFLLIFTVI 520
 Qy 121 QFHRPNYGAIPYDWCVALGCMIVFCIIMPIIMAIKIITQAKN-IFQRLISCCRPASN 179
 Db 521 QYRPITYNHQYCGMAVARIQFLMASSVICIPYALFQLCRTDGLTLQRLKNAKTRPSRD 580
 Qy 180 WGPYLEQHRGERYKDMVVPKKEAGHEI 206
 Db 581 WGPALLEHRTGRVAPTTTSPEDGEFV 607

RESULT 8

dopamine transporter protein - rat
 S20346
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
 C:Accession: S20346
 R:Giros, B.; el Mestikawy, S.; Bertrand, L.; Caron, M.G.
 FEBS Lett. 295, 149-154, 1991
 A:Title: Cloning and functional characterization of a cocaine-sensitive dopamine tran
 A:Reference number: S20346; MUID:92111727; PMID:1765147
 A:Accession: S20346
 A:Molecule type: mRNA
 A:Residues: 1-619 <GIR>
 A:Cross-references: GB:S76145; NID:g243202; PIDN:AAB21099.1; PID:g243203
 C:Superfamily: gamma-aminobutyric acid transporter
 C:Keywords: transmembrane protein

Query Match

Best Local Similarity 27.3%; Score 327; DB 2; Length 619;
 Matches 64; Conservative 32; Mismatches 74; Indels 4; Gaps 2;

Qy 3 ETTITTTIODELPK--VVKMKRVPTLGGCLVFLGLVCTQAGIYVWHLIDHFCAGWGI 62

```
Db      427 ESVITGAVDVF-QILNRHRELFLLIGLVATFLLSFCVTNGSVLYFTLDDHFAAGSTILE 485
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Oy      63 AALLEVGIIWYGNRFIEDTEBMIGAKRWTFMTMRACVITPILLIAFIWSLVQF 122
        ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      486 GVLEIEAGVAMFGVGOFSDDIKOMGORP---NIYWRICMLVSPCELLVVVVSIVTF 542
        :||::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy      123 HRPBYGAIPEDBGVALGCMIYEFTIWIPIAIIKIIIOAKNIFORLISCPR 176
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      543 RPRHYGVITPPDMANALGMILATSSAMMPITYATYTKFCGLPSDFEKLAYATTP 596
```


A:Title: Phylogenetic conservation of 4-aminobutyric acid (GABA) transporter isoforms: d

A:Reference number: S63539; MUID:96163462; PMID:8575437

A:Accession: S63539

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-622 <GUT>

A:Cross-references: EMBL:X87170; NID:g1171629; PIDN:CAA60635.1; PID:g1171630

C:Superfamily: gamma-aminobutyric acid transporter

Query Match 25.1%; Score 300.5; DB 2; Length 622;
Best Local Similarity 34.3%; Pred. No: 1.6e-21;
Matches 62; Conservative 40; Mismatches 72; Indels 7; Gaps 4;

```
QY 3 EITTTIDLPKVMK--MRVPTLGGCVLFLGLVCTQAGIYVWHLIDHFC-AGWG 59
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 404 ESIYFAIDMYPPIFRGRREDLIFVIALASYLGLVWYIEGMYTFQLFDAYASSGMC 463
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 60 ILIAITLVLGIWYIGNRFIEDTEMIGAKRWIFWLMWRACWFIPIILIAFTWSL 119
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 464 LLEVALFEICIGWYVGNRFYDNIEDMGYPFVLKW--CWIFITPGICAAIFIFI 520
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 120 VOFHPPNYGAI-PYPMGYALGCMIVFCIIWIPIMALKIIOAKNIFORLISCCRPAS 178
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 521 VRYQPLKYNVYVYPPMGYALGWAIALSSNICIPLGIFIKMNSTECTPLEKIKLITPSA 580
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 179 N 179
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 581 D 581
```

Search completed: August 27, 2003, 17:37:13
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 12:09:04 : Search time 24 Seconds

(without alignments)
423.240 Million cell updates/sec

Title: US-09-940-227-83

Sequence: 1 SEETITTTTIDLFPRKMKM.....VPRKEAGHEIPTVSGSRKPE 216

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	504.5	42.1	799	1	6S6A5_RAT
2	502.5	41.9	797	1	6S6A5_HUMAN
3	361	30.1	636	1	6S6A7_HUMAN
4	359	30.0	637	1	6S6A7_RAT
5	345.5	28.8	692	1	6S6A9_HUMAN
6	344.5	28.8	692	1	6S6A9_RAT
7	339.5	28.3	633	1	6S6A9_MOUSE
8	336.5	28.1	638	1	6S6A9_BOVIN
9	327	27.3	619	1	6S6A3_MOUSE
10	327	27.3	619	1	6S6A3_HUMAN
11	319	26.6	620	1	6S6A3_HUMAN
12	315	26.3	693	1	6S6A3_BOVIN
13	308.5	25.8	598	1	6S6A1_MOUSE
14	308.5	25.8	598	1	6S6A1_HUMAN
15	300.5	25.1	598	1	6S6A1_RAT
16	299	25.0	617	1	6S6A1_MOUSE
17	296	24.7	599	1	6S6A2_HUMAN
18	293	24.5	617	1	6S6A2_MOUSE
19	288	24.0	620	1	6S6A6_BOVIN
20	286.5	23.9	627	1	6S6A6_RAT
21	286	23.7	602	1	6S6A2_BOVIN
22	283.5	23.7	602	1	6S6A2_MOUSE
23	283.5	23.7	602	1	6S6A2_HUMAN
24	280	23.4	622	1	6S6A4_MOUSE
25	278.5	23.2	632	1	6S6A4_MOUSE
26	277.5	23.2	632	1	6S6A4_HUMAN
27	273	22.8	641	1	6S6A4_MOUSE
28	269.5	22.5	641	1	6S6A4_HUMAN
29	269	22.5	641	1	6S6A4_RAT
30	264	22.0	614	1	6S6A6_HUMAN
31	260.5	21.7	620	1	6S6A6_HUMAN
32	259	21.6	614	1	6S6A6_RAT
33	256.5	21.4	630	1	6S6A4_BOVIN

34	256.5	21.4	635	1	6S6A8_BOVIN	018875 bos taurus
35	255.5	21.3	621	1	6S6A6_RAT	P31643 ratius norv
36	255	21.3	620	1	6S6A6_CANFA	000589 canis fam11
37	254.5	21.2	635	1	6S6A6_RABIT	P31661 oryctolagus
38	253.5	21.2	621	1	6S6A6_MOUSE	035316 mus musculu
39	253.5	21.2	621	1	6S6A6_MOUSE	P31651 mus musculu
40	252.5	21.1	635	1	6S6A6_MOUSE	P31651 mus musculu
41	252.5	21.1	635	1	6S6A6_HUMAN	P48029 homo sapien
42	251.5	21.0	635	1	6S6A8_RAT	P28570 ratius norv
43	250.5	20.9	630	1	6S6A4_MOUSE	060857 mus musculu
44	249.5	20.8	630	1	6S6A4_RAT	P31652 ratius norv
45	245.5	20.3	569	1	6S6A6_HUMAN	09nsd5 homo sapien

ALIGNMENTS

RESULT 1	ID	STANDARD:	PRT:	799 AA.
6S6A5_RAT	AC	P58295;		
16-OCT-2001 (Rel. 40, Created)	DT			
16-OCT-2001 (Rel. 40, Last sequence update)	DT			
15-SEP-2003 (Rel. 42, Last annotation update)	DT			
Sodium- and chloride-dependent glycine transporter 2 (GLYT2) (GLYT-2).	DE			
SIC6A5 OR GLYT2.	GN			
Rattus norvegicus (Rat).	OS			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	OX			
NCBI_Taxid=10116;	RN			
SEQUENCE FROM N.A.	RP			
TISSUE=Brain;	RC			
MEDLINE=94043045; PubMed=8226790;	RX			
Liu Q.-R., Lopez-Corcuera B., Mandiyan S., Nelson H., Nelson N.;	RA			
"Cloning and expression of a spinal cord- and brain-specific glycine	RT			
transporter with novel structural features."	RL			
J. Biol. Chem. 268:22802-22808(1993).				
- FUNCTION: Terminates the action of glycine by its high affinity	CC			
sodium-dependent reuptake into presynaptic terminals. May be	CC			
responsible for the termination of neurotransmission at	CC			
strychnine-sensitive glycinergic synapses.	CC			
- SUBCELLULAR LOCATION: Integral membrane protein.	CC			
- TISSUE SPECIFICITY: Specifically expressed in spinal cord, brain	CC			
stem, and to a lesser extent in the cerebellum.	CC			
- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)	CC			
family.	CC			

This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (see http://www.isb-sib.ch/announce/				
or send an email to license@sib-sib.ch).				

EMBL: L21672; NOT_ANNOTATED_CDS.	DR			
PIR: A48716; A48716.	DR			
InterPro: IPR001175; Na/ntran_symport.	DR			
Pfam: PF00209; SNF: 1.	DR			
PRINTS: PR00176; NANSUSPORT.	DR			
PRODOM: P0000446; Na/ntran_symport; 2.	DR			
PROSITE: PS00610; NA_NEUTROTRAN_SYM_1; 1.	DR			
PROSITE: PS00754; NA_NEUTROTRAN_SYM_2; 1.	DR			
PROSITE: PS0267; NA_NEUTROTRAN_SYM_3; 1.	DR			
Neurotransmitter transporter; Transport; Transmembrane; Glycoprotein;	KW			
Symport.	FT			
DOMAIN 1	FT			
TRANSMEM 201	FT			
TRANSMEM 202	FT			
TRANSMEM 230	FT			
TRANSMEM 273	FT			
DOMAIN 294	FT			
TRANSMEM 396	FT			
4 (POTENTIAL).	FT			

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M88111; AAA1541.1; ALT_TERM.
 DR PIR: JH0674; JH0674.
 DR InterPro: IPR000175; Na/nttran_sympoort.
 DR Pfam: PF00209; SNF; 1.
 DR PRINTS: PR00176; NAMEUSMPOR.
 DR ProDom: PD000448; Na/nttran_sympoort; 1.
 DR PROSITE: PS00610; NA_NEUROTTRAN_SYMP_1; 1.
 DR PROSITE: PS00754; NA_NEUROTTRAN_SYMP_2; 1.
 DR PROSITE: PS50267; NA_NEUROTTRAN_SYMP_3; 1.
 DR Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
 KM Symport; Amino-acid transport; Cytoplasmic (POTENTIAL).
 FT DOMAIN 1 45
 FT TRANSMEM 46 66
 FT TRANSMEM 74 93
 FT TRANSMEM 117 137
 FT DOMAIN 138 214
 FT TRANSMEM 215 233
 FT TRANSMEM 242 259
 FT TRANSMEM 295 312
 FT TRANSMEM 324 345
 FT TRANSMEM 378 397
 FT TRANSMEM 425 443
 FT TRANSMEM 459 479
 FT TRANSMEM 500 519
 FT TRANSMEM 538 556
 FT DOMAIN 557 637
 FT CARBOHYD 182
 SQ SEQUENCE 637 AA: 71090 MW: 9627E6DD5B8C9408 CRC64;
 Query Match 30.0%; Score 359; DB 1; Length 637;
 Best Local Similarity 30.4%; Pred. No. 9.3e-25;
 Matches 65; Conservative 55; Mismatches 78; Indels 16; Gaps 2;
 QY .3 ETTTITODLPFKPKMKRVPTTLCCLVFLGLVVCYQAGIYVWHLIDHFCAGWGLI 62
 DB 405 ERTVAIVDEPPIYLRKRAVSGLLVAMTGLITLDGMYVILDDVSASRGLAV 464
 QY 63 AAILELVGIIWYGNRTEDTDMGAKRMIFWLMWRACFVITPILIAIFMSIYOF 122
 DB 465 VITTCIAVTRYGIORFCDIHMMIGFPR--GLYFRACWFLSPATLALLVYSIVY 521
 QY 123 HRPNGCAIRPPDGVAGLGMCIWPCIIPIMAITKIIOAKGNIFORLISCCRPASNMG 182
 DB 522 QPSEGSYRPPMAELLGLMGLISCLMIPACMLVAVLEEGSLMERLQASRPALDWGP 581
 QY 183 VLEOHGERYKDMVVPKKEAGHEIPTVSGSRKPE 216
 DB 582 SLEENR-----TGMVAVATLAGSQSPK 602
 RESULT 5
 S6A9_HUMAN
 ID S6A9_HUMAN STANDARD; PRT; 692 AA.
 AC P48067;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Sodium- and chloride-dependent glycine transporter 1 (GLYT1) (GLYT-1).
 GN SLC6A9.
 OS Homo sapiens (Human).
 OS Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS GLYT-1A; GLYT-1B AND GLYT-1C).
 RC TISSUE=Brain;
 RX MEDLINE=94239375; PubMed=8183239;
 RA Kim K.-M., Kingsmore S.F., Han H., Yang-Feng T.L., Godinot N.,
 RA Seidlin M.F., Caron M.G., Giros B.;
 RT Cloning of the human glycine transporter type 1: molecular and
 RT pharmacological characterization of novel isoform variants and
 RT chromosomal localization of the gene in the human and mouse
 RT genomes".
 RL Mol. Pharmacol. 45:608-617(1994).
 CC -1- FUNCTION: Terminates the action of glycine by its high affinity
 CC sodium-dependent reuptake into presynaptic terminals. May play a
 CC role in regulation of glycine levels in NMDA receptor-mediated
 CC neurotransmission.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=GLYT-1C;
 CC IsoId=P48067-1; Sequence=Displayed;
 CC Name=GLYT-1A;
 CC IsoId=P48067-2; Sequence=VSP_006270;
 CC Name=GLYT-1B;
 CC IsoId=P48067-3; Sequence=VSP_006271;
 CC -1- TISSUE SPECIFICITY: ALI GLYT-1 SUBTYPES CAN BE FOUND IN BRAIN,
 CC KIDNEY, PANCREAS, LUNG, PLACENTA, LIVER. BUT GLYT-1C SUBTYPE IS
 CC ONLY FOUND IN THE BRAIN.
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: S70609; AAB30784.1; -;
 DR EMBL: S70612; AAB30785.1; -;
 DR PIR: I57956; I57956.
 DR PIR: I77912; I77912.
 DR GeneW: HGNC:11056; SLC6A9.
 DR MIM: 601019; -;
 DR GO: GO:0005887; C: integral to plasma membrane; TAS.
 DR GO: GO:0005624; C: membrane fraction; TAS.
 DR GO: GO:0015375; F: glycine:sodium symporter activity; TAS.
 DR GO: GO:0006832; P: small molecule transport; TAS.
 DR InterPro: IPR00175; Na/nttran_sympoort.
 DR Pfam: PF00209; SNF; 1.
 DR PRINTS: PR00176; NAMEUSMPOR.
 DR ProDom: PD000448; Na/nttran_sympoort; 2.
 DR PROSITE: PS00610; NA_NEUROTTRAN_SYMP_1; 1.
 DR PROSITE: PS00754; NA_NEUROTTRAN_SYMP_2; 1.
 DR PROSITE: PS50267; NA_NEUROTTRAN_SYMP_3; 1.
 DR Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
 KW Symport; Amino-acid transport; Transport; Alternative splicing;
 KM Symport; Amino-acid transport; Cytoplasmic (POTENTIAL).
 FT DOMAIN 1 94
 FT TRANSMEM 95 115
 FT TRANSMEM 123 142
 FT TRANSMEM 166 186
 FT TRANSMEM 187 273
 FT DOMAIN 274 292
 FT TRANSMEM 301 318
 FT TRANSMEM 354 371
 FT TRANSMEM 383 404
 FT TRANSMEM 437 456
 FT TRANSMEM 485 503
 FT TRANSMEM 519 539
 FT TRANSMEM 560 579
 FT TRANSMEM 598 616
 FT TRANSMEM 617 692
 FT DOMAIN 623 723
 FT CARBOHYD 226
 SQ SEQUENCE 726 AA: 71090 MW: 9627E6DD5B8C9408 CRC64;
 Query Match 30.0%; Score 359; DB 1; Length 637;
 Best Local Similarity 30.4%; Pred. No. 9.3e-25;
 Matches 65; Conservative 55; Mismatches 78; Indels 16; Gaps 2;
 QY .3 ETTTITODLPFKPKMKRVPTTLCCLVFLGLVVCYQAGIYVWHLIDHFCAGWGLI 62
 DB 405 ERTVAIVDEPPIYLRKRAVSGLLVAMTGLITLDGMYVILDDVSASRGLAV 464
 QY 63 AAILELVGIIWYGNRTEDTDMGAKRMIFWLMWRACFVITPILIAIFMSIYOF 122
 DB 465 VITTCIAVTRYGIORFCDIHMMIGFPR--GLYFRACWFLSPATLALLVYSIVY 521
 QY 123 HRPNGCAIRPPDGVAGLGMCIWPCIIPIMAITKIIOAKGNIFORLISCCRPASNMG 182
 DB 522 QPSEGSYRPPMAELLGLMGLISCLMIPACMLVAVLEEGSLMERLQASRPALDWGP 581
 QY 183 VLEOHGERYKDMVVPKKEAGHEIPTVSGSRKPE 216
 DB 582 SLEENR-----TGMVAVATLAGSQSPK 602
 RESULT 5
 S6A9_HUMAN
 ID S6A9_HUMAN STANDARD; PRT; 692 AA.
 AC P48067;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Sodium- and chloride-dependent glycine transporter 1 (GLYT1) (GLYT-1).
 GN SLC6A9.
 OS Homo sapiens (Human).
 OS Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]

OY 180 MGPVLEOHGERKMDVVPKKEACHEI 206
 DB 581 MGPALLEHRTGKRAFTTPSPEDGEV 607

RESULT 7
 S6A9 MOUSE STANDARD: PRT: 633 AA.
 ID S6A9 MOUSE
 AC P28571:
 DT 01-DEC-1992 (rel. 24, Created)
 DT 01-DEC-1992 (rel. 24, Last sequence update)
 DT 15-SEP-2003 (rel. 42, Last annotation update)
 DE Sodium- and chloride-dependent glycine transporter 1 (GLYT1) (GLYT-1).
 GN SLC6A9 OR GLYT1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORMS GLYT-1A AND GLYT-1B).
 RX TISSUE=Brain;
 RX MEDLINE=92316254; PubMed=1618338;
 RA Liu Q.-R., Nelson H., Mandiyan S., Lopez-Corcuera B., Nelson N.;
 RT "Cloning and expression of a glycine transporter from mouse brain";
 RL FEBS Lett. 305:110-114(1992).
 CC -1- FUNCTION: Terminates the action of glycine by its high affinity
 CC sodium-dependent reuptake into presynaptic terminals. May play a
 CC role in regulation of glycine levels in NMDA receptor-mediated
 CC neurotransmission.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=GLYT-1A;
 CC IsoId=P28571-1; Sequence=Displayed;
 CC Name=GLYT-1B;
 CC IsoId=P28571-2; Sequence=VSP_006272;
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X67056; CAA47440.1; -;
 DR PIR: S23151; S23151.
 DR MGD: MGI:95760; SLC6A9.
 DR InterPro: IPR000175; Na/nttran_symport.
 DR Pfam: PF00209; SNF_1
 DR PRINTS: PR00176; NANEUSMPORT.
 DR PRODOM: PD000448; Na/nttran_symport; 2.
 DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
 DR PROSITE: PS00754; NA_NEUROTRAN_SYM_2; 1.
 DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
 KW Neurotransmitter transporter; Transporter; Transmembrane; Glycoprotein;
 KW Symport; Amino-acid transport; Alternative splicing;
 FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 36 56 1 (POTENTIAL).
 FT TRANSSEM 64 83 2 (POTENTIAL).
 FT TRANSSEM 107 127 3 (POTENTIAL).
 FT DOMAIN 128 214 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 215 233 4 (POTENTIAL).
 FT TRANSSEM 242 259 5 (POTENTIAL).
 FT TRANSSEM 295 312 6 (POTENTIAL).
 FT TRANSSEM 324 345 7 (POTENTIAL).
 FT TRANSSEM 378 397 8 (POTENTIAL).
 FT TRANSSEM 426 444 9 (POTENTIAL).
 FT TRANSSEM 460 480 10 (POTENTIAL).
 FT TRANSSEM 501 520 11 (POTENTIAL).

FT TRANSSEM 539 557 12 (POTENTIAL).
 FT DOMAIN 558 633 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1 10 MVSQKAGML -> MASAGQVATPSPEQ (in isoform
 FT GLYT-1B).
 FT FTId=VSP_006272.
 SQ SEQUENCE 633 AA; 70775 MW; FD64AA561F470793 CRC64;
 Query Match 28.3%; Score 339.5; DB 1; Length 633;
 Best Local Similarity 31.9%; Pred. No. 5,1e-23;
 Matches 66; Conservative 54; Mismatches 80; Indels 7; Gaps 4;
 OY 3 ETTITTDLPK--VKKMKVPTTGCCVTLFLGVCYQAGIYVWHLIDHCAGMI 60
 DB 405 ETLVATVDEVGNEMILQK-RYVTLGVAAVGLGIPLSQAGIYLLMDNAAASFSL 463
 OY 61 LIAATLELVGIIWYVGNRFEDTEMMIGAKWIFWLMWRACWEVTPILLALFTSLV 120
 DB 464 VVISICMGSIMYIGCHNYPQDIQMLGFPPPLF--FQICWRFVSPALIFILLFTVI 520
 OY 121 QFHRPNYCAIRPPDWGVALGCMVYFCIWMPIAIKIIOAKN-IFORLISCCRPASN 179
 DB 521 QYRPITVNHQYPCGAVRIGFLMALSSVICIPLVAFQLCRTDSDTLQRLKATKFSRD 580
 OY 180 MGPVLEOHGERKMDVVPKKEACHEI 206
 DB 581 MGPALLEHRTGKRAFTTPSPEDGEV 607

RESULT 8
 S6A9 BOVIN STANDARD: PRT: 638 AA.
 ID S6A9 BOVIN
 AC Q28039; Q28040; Q29420;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 15-SEP-2003 (rel. 42, Last annotation update)
 DE Sodium- and chloride-dependent glycine transporter 1 (GLYT1) (GLYT-1).
 GN SLC6A9.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RA Jones E.M.C.;
 RU Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Terminates the action of glycine by its high affinity
 CC sodium-dependent reuptake into presynaptic terminals. May play a
 CC role in regulation of glycine levels in NMDA receptor-mediated
 CC neurotransmission (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=1; Synonyms=glYT-1B;
 CC IsoId=Q28039-1; Sequence=Displayed;
 CC Name=2; Synonyms=glYT-1A;
 CC IsoId=Q28039-2; Sequence=VSP_006268;
 CC Name=3; Synonyms=glYT-1F;
 CC IsoId=Q28039-3; Sequence=VSP_006269;
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC -----
DR EMBL: U52687; AAB01158.1; -
DR EMBL: U52688; AAB01159.1; -
DR EMBL: U52689; AAB01160.1; -
DR EMBL: U52690; AAB01161.1; -
DR InterPro: IPR000175; Na/ntran_symport.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/ntran_symport; 2.
DR PROSITE: PS00610; NA_NEUOTRAN_SYM_1; 1.
DR PROSITE: PS00754; NA_NEUOTRAN_SYM_2; 1.
DR PROSITE: PS50267; NA_NEUOTRAN_SYM_3; 1.
DR Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW Symport; Amino-acid transport; Alternative splicing.
FT DOMAIN 1 40 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 41 61 1 (POTENTIAL).
FT TRANSMEM 69 88 2 (POTENTIAL).
FT TRANSMEM 112 132 3 (POTENTIAL).
FT DOMAIN 133 219 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 220 238 4 (POTENTIAL).
FT TRANSMEM 247 264 5 (POTENTIAL).
FT TRANSMEM 300 317 6 (POTENTIAL).
FT TRANSMEM 329 350 7 (POTENTIAL).
FT TRANSMEM 383 402 8 (POTENTIAL).
FT TRANSMEM 431 449 9 (POTENTIAL).
FT TRANSMEM 465 485 10 (POTENTIAL).
FT TRANSMEM 506 525 11 (POTENTIAL).
FT TRANSMEM 544 562 12 (POTENTIAL).
FT DOMAIN 563 638 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 14 MAAAGPVAPSSLE -> MWGKAKAM (In isoform 2).
FT VARSPLIC 537 638 /Fntd-vsp_006268.
FT VARSPLIC 537 638 YSMRWYRIGLWALSSVICIPIYALFQFCRTDGLLHRLK
FT NAKPSRDMGSPALLEHTRRYATPTSPSEDDGLEVOPLPHD
FT KAOIPWVGSGSSRFQDSRI -> SSOTGIPLPTCOIAPAH
FT VPQLSGARPPKPMPSVNSVSLARAPLCSDSPERASNNL
FT (In isoform 3).
FT VARSPLIC 1 14 /Fntd-vsp_006269.
SQ SEQUENCE 638 AA; 71318 MW; 97ADDBF28EC2A43 CRC64;
Query Match 28.1%; Score 336.5; DB 1; Length 638;
Best Local Similarity 32.4%; Pred. No. 9.5e-23;
Matches 67; Conservative 51; Mismatches 82; Indels 7; Gaps 4;
OY 3 ETTTITODLPK--VMKKMRYPITGCVLFLGLVGVTAQGIYVHLIDHFCAGMCI 60
DB 410 ETLVTAIVDEGVNEMILQK-KTYVTLGVAVAGFLGIPLTSSQAGIYWLMDNTYASFSL 468
OY 61 LIAAILELVGIWIVGNGNFIEDTEMGAKRMIFWLMWRACMFVITPILIAIFMSLV 120
DB 469 VIISCIWCSIMYITIGHQNYFODIQMLGFPPLP---FQICMRVSPAIIFILIFSVI 525
OY 121 QHRPRYGAIPYRPMGVAGMCMYIFCIITWIMAIKITQAKN-ITQRLISCCRPASN 179
DB 526 QYQPTTYNOYQYPSNRVRIGLFMASSVICIPLVAFQFCRTDGLLHRLKNMKPSHD 585
OY 180 MGPYLEOHRGERYKDMVVPKKEAGHEI 206
DB 586 WGPALLEHTRRYATPTTSPSEDDGLEV 612
RESULT 9
SSA3_MOUSE STANDARD: PRT: 619 AA.
ID 56A3_MOUSE 060719: 09R112:
AC 061327: 060719: 09R112:
DT 01-NOV-1997 (Rel. 35. Created)
DT 28-FEB-2003 (Rel. 41. Last sequence update)
DT 15-SEP-2003 (Rel. 42. Last annotation update)
DE Sodium-dependent dopamine transporter (Da transporter) (DAT).

```

```

GN SLC6A3 OR DAT1 OR DAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxId=10090;
RX STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=9307154; PubMed=10375632;
RA Wu X., Gu H.H.;
RT "Molecular cloning of the mouse dopamine transporter and
RT pharmacological comparison with the human homologue.";
RL Gene 233:163-170(1999).
RN [2]
RP SEQUENCE OF 1-343 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=95364623; PubMed=7637582;
RA Donovan D.M., Vandenberg D.J., Perry M.P., Bird G.S., Ingersoll R.,
RA Nanthakumar E., Uhl G.R.;
RT "Human and mouse dopamine transporter genes: conservation of
RT 5'-flanking sequence elements and gene structures.";
RL Brain Res. Mol. Brain Res. 30:327-335(1995).
CC -1- FUNCTION: Amino transporter. "Terminates the action of dopamine by
CC its high affinity sodium-dependent reuptake into presynaptic
CC terminals (By similarity).
CC -1- SUBUNIT: Homooligomer; disulfide-linked. Interacts with PRKCAB
CC (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- MISCELLANEOUS: This protein is the target of psychomotor
CC stimulants such as amphetamines or cocaine.
CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
CC family.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF109072; AAD19643.1; -
DR EMBL: U15791; AAA86462.1; JOINED.
DR EMBL: U12313; AAA86462.1; JOINED.
DR EMBL: U16265; AAC52283.1; -
DR MGD: MGI:94862; Slc6a3.
DR GO: GO:0005329; F:dopamine transporter activity; IMP.
DR GO: GO:0042416; P:dopamine biosynthesis; IMP.
DR GO: GO:0042420; P:dopamine catabolism; IMP.
DR GO: GO:0015872; P:dopamine transport; IMP.
DR GO: GO:0042053; P:regulation of dopamine metabolism; IMP.
DR InterPro: IPR000175; Na/ntran_symport.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR PRODOM: PD000448; Na/ntran_symport; 1.
DR PROSITE: PS00610; NA_NEUOTRAN_SYM_1; 1.
DR PROSITE: PS00754; NA_NEUOTRAN_SYM_2; 1.
DR PROSITE: PS50267; NA_NEUOTRAN_SYM_3; 1.
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW Symport.
FT DOMAIN 1 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 1 (POTENTIAL).
FT TRANSMEM 96 116 2 (POTENTIAL).
FT TRANSMEM 140 160 3 (POTENTIAL).
FT DOMAIN 161 236 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 237 255 4 (POTENTIAL).
FT TRANSMEM 264 281 5 (POTENTIAL).
FT TRANSMEM 317 334 6 (POTENTIAL).
FT TRANSMEM 346 367 7 (POTENTIAL).
FT TRANSMEM 400 419 8 (POTENTIAL).
FT TRANSMEM 446 464 9 (POTENTIAL).
FT TRANSMEM 480 500 10 (POTENTIAL).
FT TRANSMEM 521 540 11 (POTENTIAL).

```

```

FT TRANSMEM 559 577 12 (POTENTIAL).
FT DOMAIN 578 619 CYTOPLASMIC (POTENTIAL).
FT DISULFID 305 305 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 181 181 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAc. . .) (POTENTIAL).
SQ SEQUENCE 619 AA: 68805 MW: 143617 BCE461E9B CRC64:

Query Match 27.3%; Score 327; DB 1; Length 619;
Best Local Similarity 36.8%; Pred. No. 6,5e-22;
Matches 64; Conservative 32; Mismatches 74; Indels 4; Gaps 2;

OY 3 EITTTTIOQLPFXKKKKRVPTTLCGLVLTGCVYQAGIYVWHLIDHFCAGWGLI 62
DB 427 ESVITGLVDF-QLLHRRHELTGLVATFLSLFCVYNGSIYFTLLDHFAGTSLIF 485
OY 63 AALIEVLGIWYGNRFEDTEMIGAKRMIFWLMRACWEVTPILLIAFIWSLVQF 122
DB 486 GVLIEALIGVAMFYGVQGFSDDIKQMTGQRP--NLYMRCKMLVSPCFLLYVVVSIYTF 542
OY 123 HRPNTGALPYDVGVALGCMVFCIWIPIMAIITIKIIOAKNIFQRLISCRP 176
DB 543 RPHYGVAYIFPDMANALGWIITATSSMAWPIYATYKFCSLPGSFREKLAYATP 596

RESULT 10
S6A3_RAT STANDARD: PRT: 619 AA.
ID S6A3_RAT
AC P23977;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sodium-dependent dopamine transporter (DA transporter) (DAT).
GN SLC6A3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92054540; PubMed=1948035;
RA Kelly J.E., Lorang D., Amara S.G.;
RT "Cloning and expression of a cocaine-sensitive rat dopamine
RT transporter."
RL Science 254:578-579(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92054539; PubMed=1948034;
RA Shimada S., Kitayama S., Lin C.-L., Patel A., Nanthakumar E.,
RA Gregor P., Kuhar M., Uhl G.;
RT "Cloning and expression of a cocaine-sensitive dopamine transporter
RT complementary DNA."
RL Science 254:576-578(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92111727; PubMed=1765147;
RA Gires B., el Mestikawy S., Bertrand L., Caron M.G.;
RT "Cloning and functional characterization of a cocaine-sensitive
RT dopamine transporter."
RL FEBS Lett. 295:149-154(1991).
RN [4]
RP MUTAGENESIS.
RX MEDLINE=92365446; PubMed=1502198;
RA Kitayama S., Shimada S., Xu H., Markham L., Donovan D.M., Uhl G.;
RT "Dopamine transporter site-directed mutations differentially alter
RT substrate transport and cocaine binding."
RL Proc. Natl. Acad. Sci. U.S.A. 89:7782-7785(1992).
CC -1- FUNCTION: Amine transporter. Terminates the action of dopamine by
CC its high affinity sodium-dependent reuptake into presynaptic
CC terminals.
CC -1- SUBUNIT: Homooligomer; disulfide-linked. Interacts with PRKCA/B
CC (by similarity).

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: This protein is the target of psychomotor
CC stimulants such as amphetamines or cocaine.
CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNP)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M80233; AAA41100.1; -
DR EMBL: M80570; AAA73143.1; -
DR EMBL: S76145; AAB21099.1; -
DR PIR: I59558; I59558.
DR PIR: S20346; S20346.
DR InterPro: IPR00175; Na/nttran_symport.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR ProDom: PD000448; Na/nttran_symport; 1.
DR PROSITE: PS00610; NA_NEUOTRAN_SYM_1; 1.
DR PROSITE: PS00734; NA_NEUOTRAN_SYM_2; 1.
DR PROSITE: PS50267; NA_NEUOTRAN_SYM_3; 1.
KW Neurotransmitter transporter; Transport; Transmembrane; Glycoprotein;
KW Symport.
FT DOMAIN 1 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 1 (POTENTIAL).
FT TRANSMEM 96 116 2 (POTENTIAL).
FT TRANSMEM 140 160 3 (POTENTIAL).
FT DOMAIN 161 236 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 237 255 4 (POTENTIAL).
FT TRANSMEM 264 281 5 (POTENTIAL).
FT TRANSMEM 317 334 6 (POTENTIAL).
FT TRANSMEM 346 367 7 (POTENTIAL).
FT TRANSMEM 400 419 8 (POTENTIAL).
FT TRANSMEM 446 464 9 (POTENTIAL).
FT TRANSMEM 480 500 10 (POTENTIAL).
FT TRANSMEM 521 540 11 (POTENTIAL).
FT TRANSMEM 559 577 12 (POTENTIAL).
FT DOMAIN 578 619 CYTOPLASMIC (POTENTIAL).
FT DISULFID 305 305 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 181 181 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CONFLICT 597 597 E -> K (IN REF. 3).
SQ SEQUENCE 619 AA: 68746 MW: 143617 BCE461E9B CRC64:

Query Match 27.3%; Score 327; DB 1; Length 619;
Best Local Similarity 36.8%; Pred. No. 6,5e-22;
Matches 64; Conservative 32; Mismatches 74; Indels 4; Gaps 2;

OY 3 EITTTTIOQLPFXKKKKRVPTTLCGLVLTGCVYQAGIYVWHLIDHFCAGWGLI 62
DB 427 ESVITGLVDF-QLLHRRHELTGLVATFLSLFCVYNGSIYFTLLDHFAGTSLIF 485
OY 63 AALIEVLGIWYGNRFEDTEMIGAKRMIFWLMRACWEVTPILLIAFIWSLVQF 122
DB 486 GVLIEALIGVAMFYGVQGFSDDIKQMTGQRP--NLYMRCKMLVSPCFLLYVVVSIYTF 542
OY 123 HRPNTGALPYDVGVALGCMVFCIWIPIMAIITIKIIOAKNIFQRLISCRP 176
DB 543 RPHYGVAYIFPDMANALGWIITATSSMAWPIYATYKFCSLPGSFREKLAYATP 596

RESULT 11
S6A3_HUMAN STANDARD: PRT: 620 AA.
ID S6A3_HUMAN
AC O01959; O14996;
DT 01-APR-1993 (Rel. 25, Created)

```

DT 01-APR-1993 (Rel. 25, last sequence update)
 DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE Sodium-dependent dopamine transporter (DA transporter) (DAT).
 GN SLC6A3 OR DAT1.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93061810; PubMed-1359373;
 RA Vandenbergh D.J., Persico A.M., Uhl G.R.;
 RT "A human dopamine transporter cDNA predicts reduced glycosylation,
 RT displays a novel repetitive element and provides racially-dimorphic
 RT Tag1 RFLPs.";
 RL Brain Res. Mol. Brain Res. 15:161-166(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE-Brain;
 RC MEDLINE-93024326; PubMed-1406597;
 RA Ginos B., el Mestikawy S., Godinot N., Zheng K., Han H.,
 RA Yang-Feng T., Caron M.G.;
 RT "Cloning, pharmacological characterization, and chromosome assignment
 RT of the human dopamine transporter.";
 RL Mol. Pharmacol. 42:383-390(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE-Brain;
 RC MEDLINE-94134051; PubMed-8302271;
 RA Pristupa Z.B., Wilson J.M., Hoffman B.J., Kish S.J., Niznik H.B.;
 RT "Pharmacological heterogeneity of the cloned and native human
 RT dopamine transporter: disassociation of [3H]WIN 35,428 and [3H]GBR
 RT 12,935 binding.";
 RL Mol. Pharmacol. 45:125-135(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97444282; PubMed-9300814;
 RA Kawara T., Kawakami H., Yamamura S.;
 RT "Structure and organization of the gene encoding human dopamine
 RT transporter.";
 RL Gene 195:11-18(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20349832; PubMed-1089531;
 RA Vandenbergh D.J., Thompson M.D., Cook E.H., Bendahou E., Nguyen T.,
 RA Krasowski M.D., Zarrabian D., Comings D., Sellers E.M., Tyndale R.F.,
 RA George S.R., O'Dowd B.F., Uhl G.R.;
 RT "Human dopamine transporter gene: coding region conservation among
 RT normal, Tourette's disorder, alcohol dependence and attention-deficit
 RT hyperactivity disorder populations.";
 RL Mol. Psych. 5:283-292(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21200925; PubMed-11304827;
 RA Greenwood T.A., Alexander M., Keck P.E., McElroy S., Sadovnick A.D.,
 RA Remick R.A., Kolos J.R.;
 RT "Evidence for linkage disequilibrium between the dopamine transporter
 RT and bipolar disorder.";
 RL Am. J. Med. Genet. 105:145-151(2001).
 RN [7]
 RP SEQUENCE OF 1-385 FROM N.A.
 RX MEDLINE-95364623; PubMed-7637582;
 RA Donovan D.M., Vandenbergh D.J., Perry M.P., Bird G.S., Ingersoll R.,
 RA Nantakumar E., Uhl G.R.;
 RT "Human and mouse dopamine transporter genes: conservation of
 RT 5'-flanking sequence elements and gene structures.";
 RL Brain Res. Mol. Brain Res. 30:327-335(1995).
 RN [8]
 RP SEQUENCE OF 164-285 FROM N.A.
 RX MEDLINE-92357778; PubMed-1353885;
 RA Bannon M.J., Poosch M.S., Xia Y., Goebel D.J., Cassin B.,
 RA Kapatos G.;
 RT "Dopamine transporter mRNA content in human substantia nigra

RT decreases precipitously with age.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7095-7099(1992).
 RN [9]
 RP INTERACTION WITH PRKCAP.
 RX MEDLINE-21243046; PubMed-11343649;
 RA Torres G.E., Yao W.-D., Mohr A.R., Quan H., Kim K.-M., Levey A.I.,
 RA Staudinger J., Caron M.G.;
 RT "Functional interaction between monoamine plasma membrane transporters
 RT and the synaptic PDZ domain-containing protein PICK1.";
 RL Neuron 30:121-134(2001).
 RN [10]
 RP INTERCHAIN DISULFIDE BOND.
 RA Hestrup H., Karlin A., Javitch J.A.;
 RT "Symmetrical homodimer of the human dopamine transporter revealed by
 RT cross-linking Cys306 at the extracellular end of TM6.";
 RL Abstr. - Soc. Neurosci. 27:1866-1866(2001).
 RN [11]
 RP VARIANT GLN-237.
 RX MEDLINE-99318093; PubMed-10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [12]
 RP ERRATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 CC -1- FUNCTION: Amine transporter. Terminates the action of dopamine by
 CC its high affinity sodium-dependent reuptake into presynaptic
 CC terminals.
 CC -1- SUBUNIT: Homooligomer; disulfide-linked. Interacts with PRKCAP.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- MISCELLANEOUS: This protein is the target of psychomotor
 CC stimulants such as amphetamines or cocaine.
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M9670; AAA35770.1; -;
 DR EMBL; M95167; AAC1720.1; -;
 DR EMBL; S46955; AAA11754.1; -;
 DR EMBL; S44626; AAB23443.1; -;
 DR EMBL; L24178; AAA19560.1; -;
 DR EMBL; AF321321; AAG33844.1; -;
 DR EMBL; AF306558; AAG33844.1; JOINED.
 DR EMBL; AF321320; AAG33844.1; JOINED.
 DR EMBL; AF306559; AAG33844.1; JOINED.
 DR EMBL; AF306560; AAG33844.1; JOINED.
 DR EMBL; AF306561; AAG33844.1; JOINED.
 DR EMBL; AF306562; AAG33844.1; JOINED.
 DR EMBL; AF306563; AAG33844.1; JOINED.
 DR EMBL; AF306564; AAG33844.1; JOINED.
 DR EMBL; AF191917; AAC50179.2; -;
 DR EMBL; D88570; BAA22511.1; -;
 DR EMBL; D88557; BAA22511.1; JOINED.
 DR EMBL; D88558; BAA22511.1; JOINED.
 DR EMBL; D88559; BAA22511.1; JOINED.
 DR EMBL; D88560; BAA22511.1; JOINED.
 DR EMBL; D88561; BAA22511.1; JOINED.
 DR EMBL; D88562; BAA22511.1; JOINED.


```

DR EMBL: D88563; BAA22511.1; JOINED.
DR EMBL: D88564; BAA22511.1; JOINED.
DR EMBL: D88565; BAA22511.1; JOINED.
DR EMBL: D88566; BAA22511.1; JOINED.
DR EMBL: D88567; BAA22511.1; JOINED.
DR EMBL: D88568; BAA22511.1; JOINED.
DR EMBL: D88569; BAA22511.1; JOINED.
DR PIR: A48980; A48980.
DR PIR: I57937; I57937.
DR PIR: I84455; I84455.
DR GeneW: HGNC:11049; SLC6A3.
DR MIM: 126455; -.
DR GO: GO:0005737; Cytoplasm; TAS.
DR GO: GO:0005329; Dopamine transporter activity; TAS.
DR GO: GO:0006832; Small molecule transporter; TAS.
DR InterPro: IPR000175; Na/ntan_symport.
DR Pfam: PF00209; SNF; 1.
DR PRINTS: PR00176; NANEUSMPORT.
DR ProDom: PD000448; Na/ntan_symport; 1.
DR PROSITE: PS00610; NA_NEUOTRAN_SYM_1; 1.
DR PROSITE: PS00754; NA_NEUOTRAN_SYM_2; 1.
DR PROSITE: PS0267; NA_NEUOTRAN_SYM_3; 1.
DR Neurotransmitter transporter; Transmembrane; Glycoprotein;
KW Symport; Polymorphism.
FT DOMAIN 1 68
FT TRANSMEM 69 89
FT TRANSMEM 96 116
FT TRANSMEM 140 160
FT DOMAIN 161 237
FT TRANSMEM 238 256
FT TRANSMEM 265 282
FT TRANSMEM 318 335
FT TRANSMEM 347 368
FT TRANSMEM 401 420
FT TRANSMEM 447 465
FT TRANSMEM 481 501
FT TRANSMEM 522 541
FT TRANSMEM 560 578
FT DOMAIN 579 620
FT DISULFID 306 306
FT CARBOHYD 181 181
FT CARBOHYD 188 188
FT CARBOHYD 205 205
FT VARIANT 237 237
FT CONFLICT 35 35
FT CONFLICT 345 345
FT CONFLICT 354 354
FT CONFLICT 367 367
SQ SEQUENCE 620 AA; 68494 MW; CD95009B6BA93108 CRC64;

Query Match 26.6%; Score 319; DB 1; Length 620;
Best Local Similarity 36.2%; Pred. No. 3.4e-21;
Matches 63; Conservative 32; Mismatches 75; Indels 4; Gaps 2;

OY 3 ETTTITDLPKVKMKRVPITLGGCLVFLGLVCVQAGIYWHLIDHFCAGWGIIT 62
DB 428 ESVITGLIDEF-QLHRRHRELFLEYLFLSLFCVTNGIYVFLDLHFFAGTSLIF 486
OY 63 AAILLVGIIITWYGNRIETDEMTGAKKRWIIFMLMRACGFVTPILLIAFIIMSLVOF 122
DB 487 GVLIAIVAMFYGGVSDIDQNTGQRP---SLYRLCKKLVSPCELLFVVVSVITF 543
OY 123 HRPNGAIPYDGMGVALGCMIVFCIIWIPIMAIKIIQARGNIFRLISCCRP 176
DB 544 RPHHGAIYFPDMNALGCVLATSSMAVPIYAAKFKCSLPSREKLIAYAIAP 597

```

```

DT 01-AUG-1992 (Rel. 23, last sequence update)
DT 15-SEP-2003 (Rel. 42, last annotation update)
DE Sodium-dependent dopamine transporter (DA transporter) (DAT).
GN SLC6A3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=92107910; PubMed=1722321;
RX Usdin T.B., Mezey E., Chen C., Brownstein M.J., Hoffman B.J.;
RA "Cloning of the cocaine-sensitive bovine dopamine transporter.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:1168-1171(1991).
CC -1- FUNCTION: Amine transporter. Terminates the action of dopamine by
CC its high affinity sodium-dependent reuptake into presynaptic
CC terminals.
CC -1- SUBUNIT: Homooligomer; Disulfide-linked. Interacts with PRKCAP
CC (by similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: This protein is the target of psychomotor
CC stimulants such as amphetamines or cocaine.
CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNP)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M80234; -; NOT_ANNOTATED_CDS.
CC PIR: A41617; A41617.
CC InterPro: IPR000175; Na/ntan_symport.
CC Pfam: PF00209; SNF; 1.
CC PRINTS: PR00176; NANEUSMPORT.
CC ProDom: PD000448; Na/ntan_symport; 1.
CC PROSITE: PS00610; NA_NEUOTRAN_SYM_1; 1.
CC PROSITE: PS00754; NA_NEUOTRAN_SYM_2; 1.
CC PROSITE: PS0267; NA_NEUOTRAN_SYM_3; 1.
DR Neurotransmitter transporter; Transmembrane; Glycoprotein;
KW Symport.
FT DOMAIN 1 68
FT TRANSMEM 69 89
FT TRANSMEM 97 116
FT TRANSMEM 140 160
FT TRANSMEM 161 234
FT DOMAIN 235 253
FT TRANSMEM 262 279
FT TRANSMEM 315 332
FT TRANSMEM 344 365
FT TRANSMEM 398 417
FT TRANSMEM 444 462
FT TRANSMEM 478 498
FT TRANSMEM 519 538
FT TRANSMEM 557 575
FT DOMAIN 576 633
FT DISULFID 303 303
FT CARBOHYD 181 181
FT CARBOHYD 196 196
FT CARBOHYD 202 202
SQ SEQUENCE 693 AA; 75691 MW; 0022A6FD06098660 CRC64;

Query Match 26.3%; Score 315; DB 1; Length 693;
Best Local Similarity 31.6%; Pred. No. 8.5e-21;
Matches 71; Conservative 38; Mismatches 98; Indels 18; Gaps 4;

OY 3 ETTTITDLPKVKMKRVPITLGGCLVFLGLVCVQAGIYWHLIDHFCAGWGIIT 62

```



```

Db 425 ESVTGLADEF-QLLHRRLEFLLVLALELLSCFVYNGGIYVFLLDHFAAGSTILF 483
Qy 63 AAILLVGIITWYGNRFIEDTEMIGAKRWIFMLWRACWFVITPILLAIPIFMSLVQF 122
Db 484 GVLMEVIGVAMFVGMQSDSDIKMTGRP---SLYRLCKEYSPCLLVVVVSYATF 540
Qy 123 HRPNYGAIPYDGMVGLWCMVIFCIWIPIMAIKIIQAKNIFORLISCCRPASMWG- 181
Db 541 RPHYGAIVPEPMATALGMALNASSMSVPIYAAVYKLCSLPGSSREKLVAITPETHGR 600
Qy 182 -----PYLEQ--HREGRYKDMVVPKKEAGHEIPYVSGSR 213
Db 601 VDSCGAPVHAAPLAKGVGRWRKRKSCWVPSRGPGRGPPRSPR 645

RESULT 13
S6AL_MOUSE
ID S6AL_MOUSE STANDARD: PRT: 598 AA.
AC P31646:
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sodium- and chloride-dependent GABA transporter 1.
GN SLC6A1 OR GABT1 OR GAT-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333531; PubMed=1631167;
RA Liu Q.-R., Mandlyan S., Nelson H., Nelson N.;
RL Proc. Natl. Acad. Sci. U.S.A. 89:6639-6643(1992).
CC -1- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY
SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92377; -; NOT_ANNOTATED_CDS.
DR EMBL; M92378; -; NOT_ANNOTATED_CDS.
DR PIR; F46027; F46027.
DR MGD; MGI:95627; Gabl1.
DR InterPro; IPR000175; Na/nttran_symport.
DR Pfam; PF00209; SNF; 1.
DR PRINTS; PR00176; NAMEUSMPORT.
DR PRODOM; PD000448; Na/nttran_symport; 1.
DR PROSITE; PS00610; NA_NEUROTRAN_SYM_1; 1.
DR PROSITE; PS00754; NA_NEUROTRAN_SYM_2; 1.
DR PROSITE; PS50267; NA_NEUROTRAN_SYM_3; 1.
DR Neurotransmitter transporter; Transport; Transmembrane; Glycoprotein;
KW Domain; Multigene family.
FT TRANSMEM 53 73 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 100 1 (POTENTIAL).
FT TRANSMEM 124 144 2 (POTENTIAL).
FT TRANSMEM 145 211 3 (POTENTIAL).
FT TRANSMEM 212 229 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 238 255 4 (POTENTIAL).
FT TRANSMEM 291 308 5 (POTENTIAL).
FT TRANSMEM 320 341 6 (POTENTIAL).

```

```

FT TRANSMEM 374 393 8 (POTENTIAL).
FT TRANSMEM 421 439 9 (POTENTIAL).
FT TRANSMEM 456 476 10 (POTENTIAL).
FT TRANSMEM 497 516 11 (POTENTIAL).
FT TRANSMEM 535 553 12 (POTENTIAL).
FT DOMAIN 554 598 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 598 AA; 66841 MW; 893C3575DB99BD00 CMC64;

Query Match 25.8%; Score 308.5; DB 1; Length 598;
Best Local Similarity 31.2%; Pred. No. 2.8e-20;
Matches 67; Conservative 38; Mismatches 89; Indels 21; Gaps 4;

Qy 3 EITITTIQDLPFKMKKMRPITIGCCLVFLGLVCYTOAGIYWHILIDHFCAGKIL 61
Db 401 EGIITLVDEYPLRRRLRRELEFAVSVLSLITGIGVYFRLFDYSSAGMSLL 460
Qy 62 IAILLVGIITWYGNRFIEDTEMIGAKRWIFMLWRACWFVITPILLAIPIFMSLVQ 121
Db 461 FLVFECEVSISMFQVNRFDYNDIDEMGSRCT---WKKLCSFTPIYAGVLFSAVO 517
Qy 122 FHRPNYGAIPYDGMVGLWCMVIFCIWIPIMAIKIIQAKNIFORLISCCRPASMWG 181
Db 518 MTPPLTMSYVFPKMGQGVGMALMSVLIPIGYAAVFLTKSLKORLQVMIQPS---- 573
Qy 182 PYLEQHGERYKDMVVPKKEAGHEIPYVSGSRKE 216
Db 574 -----EDIVRP--ENGPEOPAGSSASKE 595

RESULT 14
S6AL_MOUSE
ID S6AL_MOUSE STANDARD: PRT: 599 AA.
AC P23978:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sodium- and chloride-dependent GABA transporter 1.
GN SLC6A1 OR GABT1 OR GAT-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX TISSUE-BRAIN:
RA MEDLINE=90378307; PubMed=1975955;
RA Gustella J., Nelson N., Nelson H., Czyzyk L., Keynan S.,
RA Miedel M.C., Davidson N., Lester H.A., Kanner B.I.;
RA "Cloning and expression of a rat brain GABA transporter.";
RL Science 249:1303-1306(1990).
RN [2]
RP TOPOLOGY.
RX MEDLINE=97313439; PubMed=9169433;
RA Clark J.A.;
RA "Analysis of the transmembrane topology and membrane assembly of the
GAT-1 gamma-aminobutyric acid transporter.";
RL J. Biol. Chem. 272:14695-14704(1997).
CC -1- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY
SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M59742; AAA63487.1; -.
DR PIR; A35918; ACRGT.
DR InterPro; IP000175; Na/nttran_symport.
DR Pfam; PF00209; SNE; 1.
DR PRINTS; PR00176; NAMEUSMPORT.
DR PRODOM; PD000448; Na/nttran_symport; 1.
DR PROSITE; PS00610; NA_NEUOTRAN_SYM_1; 1.
DR PROSITE; PS00754; NA_NEUOTRAN_SYM_2; 1.
DR PROSITE; PS50267; NA_NEUOTRAN_SYM_3; 1.
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW Symport; Multigene family.
FT DOMAIN 1 52 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 53 73 1 (PROBABLE).
FT TRANSMEM 74 80 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 81 100 2 (PROBABLE).
FT TRANSMEM 101 123 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 124 144 3 (PROBABLE).
FT TRANSMEM 145 211 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 212 230 4 (PROBABLE).
FT TRANSMEM 231 238 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 239 256 5 (PROBABLE).
FT TRANSMEM 257 291 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 292 309 6 (PROBABLE).
FT TRANSMEM 310 320 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 321 342 7 (PROBABLE).
FT TRANSMEM 343 374 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 375 394 8 (PROBABLE).
FT TRANSMEM 395 421 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 422 440 9 (PROBABLE).
FT TRANSMEM 441 456 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 457 477 10 (PROBABLE).
FT TRANSMEM 478 497 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 498 517 11 (PROBABLE).
FT TRANSMEM 518 535 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 536 554 12 (PROBABLE).
FT TRANSMEM 555 599 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 599 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT TRANSMEM 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT TRANSMEM 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT TRANSMEM 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 599 AA; 67001 MW; 4FEF85092DC1D045 CRC64;

Query Match 25.8%; Score 308.5; DB 1; Length 599;
Best Local Similarity 31.2%; Pred. No.2.8e-20;
Matches 67; Conservative 38; Mismatches 89; Indels 21; Gaps 4.

OY 3 ETITTTIDLPKVKRKNRVITLIGCCVLTLEIGVCTOAGIYVWHLIDHFCA-GMGIL 61
DB 402 EGFITALVDEKPRLLRNRELEFIAVAVCTSVLTIGISNTGGIYFKLFDYSSAGMGL 461
OY 62 IAAIIELVGIITWYGGNFIETDEMIGAKRWIEMWRACWVITPLILIAIFWLSLVQ 121
DB 462 FLVFECCASIMFWYGVNRFYVNIQEMWGSRCI--MWKLSWSEFTPIIVAGVLFESAQ 518
OY 122 FHRPNRYAIPPDGMAVGWGCIVYCIWIPIIMATIKITLQAGNIFORLISCRFASWNG 181
DB 519 MTPPLTMSGYVPPKMGQGVGMIALMSVLIIEYNAVMPLITLKGSLKQLQWMIQPS----- 574
OY 182 PYLEQHRGERRYKDMVVPKKEAGHELPTVSGSKRPE 216
DB 575 -----EDIVRP--ENQPEQPQAGSSASKE 596

RESULT 15
S6AL MUSCO STANDARD; PRT; 598 AA.
AC P48057;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sodium- and chloride-dependent GABA transporter 1.

```

OS SL60A1 OR GATB1 OR GAT-1.
 GN Mus cookii (cook's mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10098;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=95049273; PubMed=7960586;
 RA Ruiz M., Egal H., Sathya V.P., Qian X.J., Sarker H.K.;
 RT "Cloning, expression, and localization of a mouse retinal gamma-
 RL aminobutyric acid transporter.";
 RL Invest. Ophthalmol. Vis. Sci. 35:4039-4048(1994).
 CC -1- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY
 CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
 CC STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
 CC -1- SIMILARITY: Belongs to the sodium:neurotransmitter symporter (SNF)
 CC family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL, L32178; AAA37663.1; -.
 DR MGD; MGI:95627; Gabcl.
 DR InterPro; IPR00175; Na/nttrau_symport.
 DR Pfam; PF00209; SNF; 1.
 DR PRINTS; PR00176; NANEUSMPORT.
 DR Prodom; PD00048; Na/nttrau_symport; 1.
 DR PROSITE; PS00610; NA_NEUOTRAN_SYM_1; 1.
 DR PROSITE; PS00754; NA_NEUOTRAN_SYM_2; 1.
 DR PROSITE; PS50267; NA_NEUOTRAN_SYM_3; 1.
 KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
 KW Symport; Multigene family.
 FT DOMAIN 1 52 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 53 73 1 (POTENTIAL).
 FT TRANSMEM 81 100 2 (POTENTIAL).
 FT TRANSMEM 124 144 3 (POTENTIAL).
 FT DOMAIN 145 211 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 212 229 4 (POTENTIAL).
 FT TRANSMEM 238 255 5 (POTENTIAL).
 FT TRANSMEM 291 308 6 (POTENTIAL).
 FT TRANSMEM 320 341 7 (POTENTIAL).
 FT TRANSMEM 374 393 8 (POTENTIAL).
 FT TRANSMEM 421 439 9 (POTENTIAL).
 FT TRANSMEM 456 476 10 (POTENTIAL).
 FT TRANSMEM 497 516 11 (POTENTIAL).
 FT TRANSMEM 535 553 12 (POTENTIAL).
 FT DOMAIN 554 598 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 598 AA: 66783 MW: 788EC99402EC71C3 CRC64;

Wed Aug 27 18:31:21 2003

us-09-940-227-83.rsp

Page 13

[illegible]

Search completed: August 27, 2003, 17:34:41
Job time : 25 secs

THIS PAGE BLANK (USPTO)

Q7 564 GAGCTGCTGGTTGTATACGCTATCCTTTGATTGCATATTTATCTGGTCAATTGG 633

Db 2033 AAGTCTGCTGGGCAATTTTAAACCCACACATTTTAACCTTTACCTTTGCTTACGCTTT 2092
QY 624 TGCATTTTCATAGACCTTAATATGCGGCAATTCATACCTGAGGAGAGTTGCTTTAG 683
Db 2093 AACATGSGGAACCCATGACATGATGCTCTTACCGCTATCCATGATGCTGCTGCTG 2152
QY 684 GCTGTGTATGATTTTCTTCTGATTTATTTGATTTGATTTGATTTGATTTGATTTGATTT 743
Db 2153 GATGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2212
QY 744 TTCAGGCTAAAGGAACATCTTTCAACGCTTATAGTTGCTGACAGACGCTTCTACT 803
Db 2213 ATCTGGCCCCCGAAGATTTATGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2272
QY 804 GGGGTCCATACCTGGAACACATCTGCTGGGGAAGATATTAAGACATGCTAGTTC 858
Db 2273 GGGGCCCATTTCTTACCTCAACACCGCGGAGCGCTTACAGAAACATGATCGACCC 2327

RESULT 2
US-09-191-468-121
; Sequence 121, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 121
; LENGTH: 2394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2391)
; OTHER INFORMATION: Seq50 Consensus; nt 1-2394
US-09-191-468-121

Query Match 13.7%; Score 149.4; DB 4; Length 2394;
Best Local Similarity 54.6%; Pred. No. 2,1e-30;
Matches 325; Conservative 0; Mismatches 261; Indels 9; Gaps 1;

QY 264 CCTCTGAGAAACGATCACACAAATTCAGATTATTTCCCAAGATGAGAGAAA 323
Db 1742 CCTCATGAGAACATGACATGACCTCATCTCAGACAGATTCCCAAGTACGACGAC 1801
QY 324 TGAGGCTTCCATTAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383
Db 1802 ACAAGCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861
QY 384 TGACACAGCTGGAATTTACTGGGTCATCTGATTTGACACACTGCTGCTGCTGCTGCT 443
Db 1862 TCACCTAGGAGTATTAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1921
QY 444 TTTTAAATGAGTATTAATGAGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 503
Db 1922 TTTTAAATGAGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1981
QY 504 TCATTTGAGATACAGAAATGATTTGAGCAAGAGGTTGATTTGCTGCTGCTGCTGCTG 563
Db 1982 TCTGTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2032
QY 564 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 623
Db 2033 AAGTCTGCTGGGCAATTTTAAACCCACACATTTTAACCTTTACCTTTGCTTACGCTTT 2092
QY 624 TGCATTTTCATAGACCTTAATATGCGGCAATTCATACCTGAGGAGAGTTGCTTTAG 683

Db 2093 AACATGSGGAACCCATGACATGATGCTCTTACCGCTATCCATGATGCTGCTGCTG 2152
QY 684 GCTGTGTATGATTTTCTTCTGATTTATTTGATTTGATTTGATTTGATTTGATTTGATTT 743
Db 2153 GATGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2212
QY 744 TTCAGGCTAAAGGAACATCTTTCAACGCTTATAGTTGCTGACAGACGCTTCTACT 803
Db 2213 ATCTGGCCCCCGAAGATTTATGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2272
QY 804 GGGGTCCATACCTGGAACACATCTGCTGGGGAAGATATTAAGACATGCTAGTTC 858
Db 2273 GGGGCCCATTTCTTACCTCAACACCGCGGAGCGCTTACAGAAACATGATCGACCC 2327

RESULT 3
US-08-700-013B-14
; Sequence 14, Application US/08700013B
; Patent No. 5919653
; GENERAL INFORMATION:
; APPLICANT: Albert, Vivian R.
; APPLICANT: Kowalski, Leslie R.Z.
; APPLICANT: Borden, Laurence A.
; APPLICANT: McKelvy, Jeffrey F.
; TITLE OF INVENTION: Human Glycine Transporter
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,013B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: 317743-108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-700-013B-14

Query Match 13.5%; Score 147.8; DB 2; Length 840;
Best Local Similarity 54.7%; Pred. No. 3.9e-30;
Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;

QY 272 GAAAGCATCAACAACAATTAATTTATTTCCCAAGATGATGAGAAATGAGGTT 331
Db 211 GAGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 270
QY 332 CCCATTAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391
Db 271 GTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 330

```

Oy 392 GCTGGAATTTACTGGGTTTCATCTGATTCAGCCATTCCTGCTGATGGGCAATTTAAAT 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 GGTGGAATTTACAGTTTACAGTTTGTGACACATTCGTGCTCTATGCCCTTGTCTATC 390
Oy 452 GCAGCTACTGAGACTAGTTGGAATTCATCTGATTTATGAGGGAACAGATTCATGAG 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 ATTGCCATTTTGAAGCTGTGGGATCTCTTATGATGATGCTGCAAAAGATTCGTGA 450
Oy 512 GATACAGAAATGATGATGAGCAAAAGAGTGATATTCCTGCTATGCTGAGAGCTTGC 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 572 TGGTTTGAATTAAGCCATCTCTTTCATTCATTAATTAATTAATTAATTAATTAAT 631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TGGGCAATTTTAACCCCAACCATTTTAACCTTATCCCTTTCAGCTTTTACAGTGG 561
Oy 632 CATGACCTAATTTATGGCGCAATTCATACCTCAGTGGGAGTGGCTTTAGGCTGTGT 691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 GAGCCATGACCTATGGCTTTCACCTATCCTAAGTGTGCTGATGCTGAGATGGCTA 621
Oy 692 ATGATTTTCTGATTAATTTGATTCATTAATTAATTAATTAATTAATTAATTAAT 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 ATGCTCCGCTGTCCGTCATCTGATCCCAATTTATGTTGATTAATAATGATCTGACC 681
Oy 752 AAAGCAACATCTTTCAACGCCCTTAATTAATTAATTAATTAATTAATTAATTAAT 811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 CCTGGAAGATTTATGAGAGCTGAAGTTGCTGTGCTGCGCAACGCCGAGTGGGCCA 741
Oy 812 TACTGGAACACATCTGAGGGAAGATTAATAACATGATGCTCC 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 742 TTCCTACCTCAACACCGCGGGAGCGTTACAAAGACATGATGACCC 788

```

RESULT 4

```

US-08-700-013B-16
: Sequence 16 Application US/08700013B
: Patent No. 5919653
: GENERAL INFORMATION:
: APPLICANT: Albert, Vivian R.
: APPLICANT: Kowalski, Leslie R.Z.
: APPLICANT: Borden, Laurence A.
: APPLICANT: McKelvy, Jeffrey F.
: TITLE OF INVENTION: Human Glycine Transporter
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: STREET: 997 Lenox Drive, Building 3, Suite 210
: CITY: Lawrenceville
: STATE: NJ
: COUNTRY: USA
: ZIP: 08543
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/700,013B
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Bloom, Allen
: REGISTRATION NUMBER: 29,135
: REFERENCE/DOCKET NUMBER: 317743-108
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 609-520-3214
: TELEFAX: 609-520-3259
: TELEX:
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 840 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-700-013B-16

```

```

Query Match 13.5%; Score 147.8; DB 2; Length 840;
Best Local Similarity 54.7%; Pred. No. 3.9e-30;
Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;

```

```

Oy 272 GAACGATACACACAAATTCAGATTTATTCGCCAAGATGATGAAGAAATGAGGTT 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 GAGACCAATTAAGACCTCCTCATCTCAGAGATTTCCCAAGTACACGACACACAGCA 270
Oy 332 CCATTAATTTGGGCTCTGCTGTTGTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 GTGTTTACTCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330
Oy 392 GCTGGAATTTACTGGGTTTCATCTGATTCAGCCATTCCTGCTGATGGGCAATTTAA 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 GGTGGAATTTACAGTTTACAGTTTGTGACACCTATGCTGCTGCTGCTGCTGCTGCTG 390
Oy 452 GCAGCTACTGAGACTAGTTGGAATTCATCTGATTTATGAGGGAACAGATTCATGAG 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 ATTGCCATTTTGAAGCTGTGGGATCTCTTATGATGATGCTGCAAAAGATTCGTGA 450
Oy 512 GATACAGAAATGATGATGAGCAAAAGAGTGATATTCCTGCTATGCTGAGAGCTTGC 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 572 TGGTTTGAATTAAGCCATCTCTTTCATTCATTAATTAATTAATTAATTAATTAAT 631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TGGGCAATTTTAACCCCAACCATTTTAACCTTATCCCTTTCAGCTTTTACAGTGG 561
Oy 632 CATGACCTAATTTATGGCGCAATTCATACCTCAGTGGGAGTGGCTTTAGGCTGTGT 691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 GAGCCATGACCTATGGCTTTCACCTATCCTAAGTGTGCTGATGCTGAGATGGCTA 621
Oy 692 ATGATTTTCTGATTAATTTGATTCATTAATTAATTAATTAATTAATTAATTAAT 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 ATGCTCCGCTGTCCGTCATCTGATCCCAATTTATGTTGATTAATAATGATCTGACC 681
Oy 752 AAAGCAACATCTTTCAACGCCCTTAATTAATTAATTAATTAATTAATTAATTAAT 811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 CCTGGAAGATTTATGAGAGCTGAAGTTGCTGTGCTGCGCAACGCCGAGTGGGCCA 741
Oy 812 TACTGGAACACATCTGAGGGAAGATTAATAACATGATGCTCC 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 742 TTCCTACCTCAACACCGCGGGAGCGTTACAAAGACATGATGACCC 788

```

RESULT 5

```

US-09-191-468-123
: Sequence 123 Application US/09191468A
: Patent No. 6416975
: GENERAL INFORMATION:
: APPLICANT: Gallagher, Michael J.
: APPLICANT: Burgess, Loyd R.
: APPLICANT: Brunden, Kurt R.
: TITLE OF INVENTION: Human Glycine Transporter Type 2
: FILE REFERENCE: 1231U0S01
: CURRENT APPLICATION NUMBER: US/09/191,468A
: CURRENT FILING DATE: 1998-11-12
: NUMBER OF SEQ ID NOS: 124
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 123
: LENGTH: 2394
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(2391)
: OTHER INFORMATION: SEQ ID NO:26 [W098/07854 (PCT/US97/14637)] All611x

```

OTHER INFORMATION: Sequence: nt 1-2394; nt 304 may be G; nt 371 may
 OTHER INFORMATION: be T; nt 836 may be A; nt 1116 may be G; nt 1831
 OTHER INFORMATION: may be G; nt 2382 may be A or T; nt 2385 may be G;
 FEATURE:
 OTHER INFORMATION: nt 2388 may be A
 US-09-191-468-123

Query Match 13.5%; Score 147.8; DB 4; Length 2394;
 Best Local Similarity 54.7%; Pred. No. 5.7e-30;
 Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;

272 GAACGATCAACAACAATTTCAAGATTTATTTCCCAAGTGAAGAAATGAGGTT 331
 1750 GAGACCATAGTACCTCCATCTCAGACAGAGTTTCCCAAGTACCTAGCAGACAGCA 1809
 332 CCCATACCTTGGGCTGCTGCTGTTGTTTCTCTCTGCTGCTGCTGCTGCTGCTG 391
 1810 GTGTTACTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1869
 392 GCTGGAATTTACGGTTCATCTGATTCACCTCTGCTGCTGCTGCTGCTGCTGCTG 451
 1870 GGTGGAATTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1929
 452 GCAGTATACCTGAGCTAGTTCGATTCATCTGATTCATTCGATTCATTCGATTCAT 511
 1930 ATTCGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1989
 512 GATACAGAAATGATGATTCGAGCAAGAGTGGATTTCTGCTGCTGCTGCTGCTGCTG 571
 1990 GATATAGATGATGATTCGATTCATTCGATTCATTCGATTCATTCGATTCATTCG 2040
 572 TGGTTGTAATTCAGCTATCTTTGATTCGATTCGATTCGATTCGATTCGATTCGAT 631
 2041 TGGGCAATTTGTAACCCCAACATTTTAACCTTATCTTTGCTTACGCTTTTACCACTG 2100
 632 CATAGACTAATTAATGCGCAATTCATCCCTGACTGAGGAGCTTCTTACGCTGCTG 691
 2101 GAGCCCATGACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
 692 ATGATGTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751
 2161 ATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
 752 AAGGAACATCTTTCAGCTTATTAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
 2221 CCTGGAGATTTATTTGAGAGCTGGAAGTTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
 812 TACCTGACACATGCTGGGGAAGATTAATAAGACATGCTGCTGCTGCTGCTGCTGCTG 858
 2281 TTCTTAGCTCAACACCGGGGAGCGTTTACAGAACATGATGACCC 2327

RESULT 6

US-08-700-013B-18
 Sequence 18, Application US/08700013B
 Patent No. 5918653

GENERAL INFORMATION:

APPLICANT: Albert, Vivian R.
 APPLICANT: Kowalski, Leslie R.Z.
 APPLICANT: Borden, Laurence A.
 APPLICANT: McKelvy, Jeffrey F.
 TITLE OF INVENTION: Human Glycine Transporter
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Dechert Price & Rhoads
 STREET: 997 Lenox Drive, Building 3, Suite 210
 CITY: Lawrenceville
 STATE: NJ
 COUNTRY: USA

ZIP: 08543

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/700,013B

FILING DATE:

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Bloom, Allen

REGISTRATION NUMBER: 29,135

REFERENCE/DOCKET NUMBER: 317743-108

TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-520-3214

TELEFAX: 609-520-3259

TELEX:

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 2397 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-700-013B-18

Query Match 13.5%; Score 147.8; DB 2; Length 2397;
 Best Local Similarity 54.7%; Pred. No. 5.7e-30;
 Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;

272 GAACGATCAACAACAATTTCAAGATTTATTTCCCAAGTGAAGAAATGAGGTT 331
 1750 GAGACCATAGTACCTCCATCTCAGACAGAGTTTCCCAAGTACCTAGCAGACAGCA 1809
 332 CCCATACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
 1810 GTGTTACTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1869
 392 GCTGGAATTTACGGTTCATCTGATTCACCTCTGCTGCTGCTGCTGCTGCTGCTG 451
 1870 GGTGGAATTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1929
 452 GCAGTATACCTGAGCTAGTTCGATTCATCTGATTCATTCGATTCATTCGATTCAT 511
 1930 ATTCGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1989
 512 GATACAGAAATGATGATTCGAGCAAGAGTGGATTTCTGCTGCTGCTGCTGCTGCTG 571
 1990 GATATAGATGATGATTCGATTCATTCGATTCATTCGATTCATTCGATTCATTCG 2040
 572 TGGTTGTAATTCAGCTATCTTTGATTCGATTCGATTCGATTCGATTCGATTCGAT 631
 2041 TGGGCAATTTGTAACCCCAACATTTTAACCTTATCTTTGCTTACGCTTTTACCACTG 2100
 632 CATAGACTAATTAATGCGCAATTCATCCCTGACTGAGGAGCTTCTTACGCTGCTG 691
 2101 GAGCCCATGACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
 692 ATGATGTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751
 2161 ATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
 752 AAGGAACATCTTTCAGCTTATTAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
 2221 CCTGGAGATTTATTTGAGAGCTGGAAGTTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
 812 TACCTGACACATGCTGGGGAAGATTAATAAGACATGCTGCTGCTGCTGCTGCTGCTG 858
 2281 TTCTTAGCTCAACACCGGGGAGCGTTTACAGAACATGATGACCC 2327

RESULT 7

US-08-700-013B-20
 Sequence 20, Application US/08700013B

Query Match	13.5%	Score 147.8	DB 2	Length 2397
Best Local Similarity	54.7%	Pred. No. 5.7e-30		
Matches 321	Conservative	0	Mismatches 257	Indels 9
			Gaps	1
QY	272	GAACGATCACAACAACAAATTCAGATTATATCCCAAGATGATGAAGAAATGAGGTT	331	
Db	1750	GAGACATAGTAGACCTCCATCTCAGACAGATTTCCCAAGTACCTACGACACACAAGCA	1805	
QY	332	CCCATAACTTTGGGCTGCTGCTGGTTTGTCTCTCTGGTCTGCTGTGTGACTCAG	391	
Db	1810	GTGTTTACTCTGGGCTGCTGCTGGTTGTTTCTTCATCATGAGGGTTTCCAAATGATCATCG	1865	
QY	392	GCTGGAAATTTCTGGGTTCAATCTGATTCAGACACTTCTGCTGGATGGGCAATTTAAT	451	
Db	1870	GGTGGAAATTCATAGTTTACGTTCTGTGACACCATATGCTGCTCTTAAGCCCTGTCTATC	1928	
QY	452	GCACCTATACAGAGCTAGTTGGAAATCATCTGGATTTATGGAGGAACAGATTCAATGAG	511	
Db	1930	ATTGCCATTTTGGAGCTCGTGGGATCTCTTATGTATGGCTTGCAGAAAGATTCGTGAA	1989	
QY	512	GATACAGAAATGATGATTTGGAGCAAGAGGTGATATTTCTGGCTATGATGGAGAGACTTGC	571	
Db	1990	GATATTAGATGATGATTTGGATTCAGAGCTTAACATCTTC-----TGGAAAGTCTGC	2040	
QY	572	TGTTTTTAATTAACGCTATCTTTTGATTTGCAATATTTATCTGGTCAATGGTGCATATT	631	
Db	2041	TGGGCATTTTGAACCCCAACCATTTTAACCTTTATCTTCTCTCAGACTTTTACCAAGTGG	2100	
QY	632	CATGACCTAATTAATGGCGCAATTCATACCTTACATGGGAGTGTGCTTATAGCTGTGTGT	691	

```

RESULT 8
US-09-182-728A-1
: Sequence 1, Application US/09182728A
: Patent No. 6238883
: GENERAL INFORMATION:
: APPLICANT: BROWN, ANTHONY
: APPLICANT: CHAPMAN, CONRAD GERALD
: APPLICANT: GLOGER, ISRAEL SIMON
: APPLICANT: EVANS, JOANNE RACHEL
: APPLICANT: CAIRNS, WILLIAM
: APPLICANT: HERDON, HUGH
: TITLE OF INVENTION: NOVEL COMPOUNDS
: FILE REFERENCE: GP-30176
: CURRENT APPLICATION NUMBER: US/09/182.728A
: CURRENT FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: 9818890.7
: PRIOR FILING DATE: 1998-08-28
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2863
: TYPE: DNA
: ORGANISM: HOMO SAPIENS
US-09-182-728A-1

```

Query Match	Best Local Similarity	13.5%	Score 147.8	DB 3	Length 2863
Matches 321	Conservative	0	Mismatches 257	Indels 9	Gaps 1
QY 272	GAAACGATCAACACACAAATTCAGATTATATCCCAAGATGATGAAGAAATGAGGCTT	331			
Db 2005	GAGACCATATGACACCTCATCTCAGACGATTTCCCAAGTACTAGCCACACACAGCA	2064			
QY 332	CCCATTAACCTTTGGGCTGCTGCTTGGTTTGTTCCTTGCTCGTCTGTGACTCAG	391			
Db 2065	GTCGTTACTCTGGGCTGCTGCATTTGTTCTTTCATCTATGGGTTTTCCAAATGATCACTCAG	2124			
QY 392	GCTGGATTTACTGCGGTTCACTGATGACACACTGCTGCTGATGGGACATTTTAAT	451			
Db 2125	GGTGGATTTACATGTTTTCAGCTGTGTGACACTATGCTGCTCTATGCCCTTGTCATC	2184			
QY 452	GCAAGCTAATCTGAGCTAGTTGSAATCATCTGATATTTATGGAGGAAACATTCATTGAC	511			
Db 2185	ATTGACCAATTTTATAGCTCGGGGAGATCTCTTATGTGATGGCTTGCAAAATCTGTGAA	2244			
QY 512	GATACACGAATGATGATTTGGAGCAAGAGGTGATATCTCGGCTATGTGGAGAAAGCTTGC	571			
Db 2245	GATATGACATGATGATTTGGATTTCCAGCCTTAACACTCTC-----TGGAAAGCTGCG	2295			
QY 572	TGGTTTGTAAATACGGCTATCTTTTGTATGCAATATTTATCTGTCATTTGGGCAATTT	631			
Db 2296	TGGGCAATTTGTAAACCCCAACACATTTTAACCTTATCTCTTTCGTTACAGCTTTTACCACTGG	2355			
QY 632	CATAGACCTAATATATGGCGCAATTCATACCTGCACTGGGAGTTGCTTTAGGCTGTGT	691			
Db 2356	GACCCCATGACCTATGAGCTCTTAACCCCTATCTCAATCGTGCATGGTGTGCTGGATGGCA	2415			
QY 692	ATGATTTGTTTTCGACTTATTTGGATTTCCAAATATATGGCTATCTAAATTAATTCAGGCT	751			

DB 2416 ATGTCGCGCTGTTCGATCTGATCCCAATTAATGTTGATATAAATGATCATCTGACC 2475
 OY 752 AAAGAAACATCTTCAACGCCCTTAATAGTTGCTGCAGACCAAGCTTCAACGGGGTCCA 811
 DB 2476 CCTGGAAGATTATTTAGAGAGCTGAAGTTGGGCTGCTGCCACACCGGAGCTGGGGCCCA 2535
 OY 812 TACCTGGAACAACATCTGCGGGAAGATATAAAGCATGTAAGTTCC 858
 DB 2536 TTCTTAGCTCAACACCGCGGGAGCGTTACAGAACATGATGACCC 2582

RESULT 9
 US-09-795-232-1
 ; Sequence 1, Application US/09795232
 ; Patent No. 6426405

GENERAL INFORMATION:
 APPLICANT: Anthony M. Brown
 APPLICANT: Conrad Gerald Chapman
 APPLICANT: Israel Simon Gloger
 APPLICANT: Joanne Rachel Evans
 APPLICANT: William Cairns
 APPLICANT: Hugh Jonathan Herdon
 TITLE OF INVENTION: NOVEL COMPOUNDS
 FILE REFERENCE: GP-30176-D1
 CURRENT APPLICATION NUMBER: US/09/795,232
 CURRENT FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: 09/182,728
 PRIOR FILING DATE: 1998-10-29
 PRIOR APPLICATION NUMBER: 9818890.7
 PRIOR FILING DATE: 1998-08-28
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 2863
 TYPE: DNA
 ORGANISM: HOMO SAPIENS
 US-09-795-232-1

Query Match 13.5%; Score 147.8; DB 4; Length 2863;
 Best Local Similarity 54.7%; Pred. No. 6.1e-30;
 Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;

OY 272 GAACGATGCAACACAAATTCAGATTTATCCCAAGTATGAAGAAATGAGGGTT 331
 DB 2005 GAGACCATGATGACCTCCATCTCAGACAGATTCCCAAGTACCTAGCCACACAAAGCCA 2064
 OY 332 CCCATACCTTTGGGCTGCTGCTGTTGTTTCTCCTTGCTCGTGTGTGACTCAG 391
 DB 2065 GTGTTTACCTGTGGGCTGCTGATTTGTTTCTTCATCATGATGATGATCCTCAG 2124
 OY 392 GCTGGAATTTACTGGGCTTATCTGATTTGACACATCTGCTGAGGAGGCGCATTTAAT 451
 DB 2125 GGTGGAATTTACTAGTTGCTGCTGAGACCTATGCTCCATGCGCTTGATC 2184
 OY 452 GCAGCTATATCTGAGTATGATGATCATCTGATTTATGAGGAGGAACATTCATTTAG 511
 DB 2185 ATTGCCATTTTATGCTCGTGGGATCTCTATGATGATGATGATGATGATGATGATGAT 2244
 OY 512 GATACGAATATGATGATGAGCAAGAGGTGATATTTGCTGATGATGAGAGCTTGC 571
 DB 2245 GATATAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2295
 OY 572 TGGTTTATATACGCTATCTGATTTGATTTGCAATTTATCTGATGATGATGATGAT 631
 DB 2296 TGGGCAATTTTAAACCCCAACCATTTTAACTTTATCTTGGCTTGAAGCTTTTACCACTG 2355
 OY 632 CATAGACCTAATATGAGCGCAATTCATACCTGATGAGGAGTGGCTTATGAGCTGATG 691
 DB 2356 GAGCCCATGACATGCTTACCGCTATACCTGATGAGGAGTGGCTTATGAGCTGATG 2415
 OY 692 ATGATGTTTTCGATTAATTTGATTCATTAATGAGCTATGATTAATTAATTCAGGCT 751

DB 2416 ATGTCGCGCTGTTCGATCTGATCCCAATTAATGTTGATATAAATGATCATCTGACC 2475
 OY 752 AAAGAAACATCTTCAACGCCCTTAATAGTTGCTGCAGACCAAGCTTCAACGGGGTCCA 811
 DB 2476 CCTGGAAGATTATTTAGAGAGCTGAAGTTGGGCTGCTGCCACACCGGAGCTGGGGCCCA 2535
 OY 812 TACCTGGAACAACATCTGCGGGAAGATATAAAGCATGTAAGTTCC 858
 DB 2536 TTCTTAGCTCAACACCGCGGGAGCGTTACAGAACATGATGACCC 2582

RESULT 10
 US-08-700-013B-24

; Sequence 24, Application US/08700013B
 ; Patent No. 5919653

GENERAL INFORMATION:
 APPLICANT: Albert, Vivian R.
 APPLICANT: Kowalski, Leslie R.Z.
 APPLICANT: Borden, Laurence A.
 APPLICANT: McKelvy, Jeffrey F.
 TITLE OF INVENTION: Human Glycine Transporter
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dechert Price & Rhoads
 STREET: 997 Lenox Drive, Building 3, Suite 210
 CITY: Lawrenceville
 STATE: NJ
 COUNTRY: USA
 ZIP: 08543
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/700,013B
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Bloom, Allen
 REGISTRATION NUMBER: 29,135
 REFERENCE/DOCKET NUMBER: 317743-108
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 609-520-3214
 TELEFAX: 609-520-3259
 TELEX:
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 589 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-700-013B-24

Query Match 13.2%; Score 144.6; DB 2; Length 589;
 Best Local Similarity 56.4%; Pred. No. 2.4e-29;
 Matches 295; Conservative 0; Mismatches 219; Indels 9; Gaps 1;

OY 336 TAACTTGGGCTGCTGCTGTTGTTTCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 395
 DB 6 TTAATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65
 OY 396 GAATTTACGCGGCTATCTGATTTGATGACCACTTCTGCTGATGAGGCGCATTTAATTGAG 455
 DB 66 GAATTTACGCTTATGCTTGTGAGACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 125
 OY 456 CTATCTGAGAGCTAGTTGATGATCTGATTTATGAGGAGGAACAGATTCATGAGAGATA 515
 DB 126 CCAATTTTGAAGCTGCGGGGAGCTCTTATGATGATGCTTGAAGAGATTCGTGAAGATA 185

TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259

TELEX:

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 2403 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-700-013B-26

Query Match

12.5% Score 136.6; DB 2; Length 2403;
Best Local Similarity 53.5%; Pred. No. 5.5e-27;
Matches 314; Conservative 0; Mismatches 264; Indels 9; Gaps 1;

```

QY 272 GAAACGATCACAACAACATTCAGATTATTTCCCAAGTGAAGAAATGAGGCTT 331
DB 1756 GAGACCATGTGACCTCCATCTCGATGATGATTCACCAATATCTGGGACACACAGCCT 1815
QY 332 CCCATACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
DB 1816 GTGTTACACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1875
QY 392 GCTGGAATTTACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
DB 1876 GGTGGATTTACATGTTTACGTTGTCGACCTATGCTGCTGCTGCTGCTGCTGCTGCTG 1935
QY 452 GCAGCTATACGAGCTAGTGGAAATCATCTGGAATTTATGAGGGGACAGATTCATTGAG 511
DB 1936 ATTGCATATTGAGCTTGTGGACATCTCTATGTCAGCGCTTGACAGAGTTCTGTGAA 1995
QY 512 GATACAGAAATGATGATGAGCAAAAGAGTGGATTTCTGCTATGAGGAGAGCTTGC 571
DB 1996 GACATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2046
QY 572 TGGTTGTAATTACGCTATCTTGTGATTCGAAATTTATCTGCTGCTGCTGCTGCTGCTG 631
DB 2047 TGGGCGTTTGACACCGACCATTTTACGTTTATCTTGTGCTGCTGCTGCTGCTGCTG 2106
QY 632 CATAGACCTAATTTAGGCGCAATTCATACCTGAGTGGGAGTTCCTTAGCTGCTGCTG 691
DB 2107 GAGCCCATACCTATGCTCTTACCGCTACCGCTACCTGCTGCTGCTGCTGCTGCTGCTG 2166
QY 692 ATGATGTTTTCGCTATTTTGGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 751
DB 2167 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2226
QY 752 AAAGAAACATCTTTCAACGCTTTATAGTTGCTGCAGACGAGCTTCTAAGTGGGCTCA 811
DB 2227 CCGGAGATTTATTTAGAGGCTGAAGTTGTTGATGCTGCACAGCCGAGCTGGGCGCCA 2286
QY 812 TACCTGGAACAACATGCTGGGAAAGATTTAAAGACATGCTAGTTC 858
DB 2287 TTCTTAGCTCAGACCGCGGGAACGCTACAAAGATATGATGACCC 2333

```

RESULT 13
US-08-655-836-4
Sequence 4, Application US/08655836
Patent No. 5824486

GENERAL INFORMATION:

APPLICANT: Borden, Laurence A.

APPLICANT: De Vivo, Michael

APPLICANT: Yokoyama, Midori

APPLICANT: Albert, Vivian R.

TITLE OF INVENTION: GLYCINE TRANSPORTER-TRANSFECTED CELLS

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dechert Price & Rhoads

STREET: P.O. Box 5218

CITY: Princeton

STATE: NJ
COUNTRY: USA
ZIP: 08543

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patenting Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/655,836

FILING DATE: 31-MAY-1996

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Bloom, Allen

REGISTRATION NUMBER: 29135

REFERENCE/DOCKET NUMBER: 317743-105

TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-520-3214

TELEFAX: 609-520-3259

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 2817 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-655-836-4

Query Match

12.5% Score 136.6; DB 1; Length 2817;
Best Local Similarity 53.5%; Pred. No. 5.8e-27;
Matches 314; Conservative 0; Mismatches 264; Indels 9; Gaps 1;

```

QY 272 GAAACGATCACAACAACATTCAGATTATTTCCCAAGTGAAGAAATGAGGCTT 331
DB 1963 GAGACCATGTGACCTCCATCTCGATGATGATTCACCAATATCTGGGACACACAGCCT 2022
QY 332 CCCATACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
DB 2023 GTGTTACACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2082
QY 392 GCTGGAATTTACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
DB 2083 GGTGGATTTACATGTTTACGTTGTCGACCTATGCTGCTGCTGCTGCTGCTGCTGCTG 2142
QY 452 GCAGCTATACGAGCTAGTGGAAATCATCTGGAATTTATGAGGGGACAGATTCATTGAG 511
DB 2143 ATGCGCATTTTGAAGCTTGTGCAATCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 2202
QY 512 GATACAGAAATGATGATGAGCAAAAGAGTGGATTTCTGCTATGCTGCTGCTGCTGCTG 571
DB 2203 GACATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2253
QY 572 TGGTTGTAATTACGCTATCTTGTGATTCGAAATTTATCTGCTGCTGCTGCTGCTGCTG 631
DB 2254 TGGGCGTTTGACACCGACCATTTTAAAGCTTTATCTGCTGCTGCTGCTGCTGCTGCTG 2313
QY 632 CATAGACCTAATTTAGGCGCAATTCATACCTGAGTGGGAGTTCCTTAGCTGCTGCTG 691
DB 2314 GAGCCCATACCTATGCTCTTACCGCTACCGCTACCTGCTGCTGCTGCTGCTGCTGCTG 2373
QY 692 ATGATGTTTTCGCTATTTTGGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 751
DB 2374 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2433
QY 752 AAAGAAACATCTTTCAACGCTTTATAGTTGCTGCAGACGAGCTTCTAAGTGGGCTCA 811
DB 2434 CCGGAGATTTATTTAGAGGCTGAAGTTGTTGATGCTGCACAGCCGAGCTGGGCGCCA 2493
QY 812 TACCTGGAACAACATGCTGGGAAAGATTTAAAGACATGCTAGTTC 858
DB 2494 TTCTTAGCTCAGACCGCGGGAACGCTACAAAGATATGATGACCC 2540

```

```

Db      2254  TGGGGCTTTTGTCACACCGCAACATTTTAAACGTTTATCTTTGCTTCAGCTGTATCATGATGG 2313
QY      632  CATAGACCTTAATTATGCGCAATTCATACCCCTGATGGGAGTGTCTTTAGGCTGGCT 691
Db      2314  GAGCCCATGACCTTAAGGCTCTTACCGCTACCCCTAAGCTGTCATAGTCTTGATGGCTG 2373
QY      692  ATGATGTGTTTCTGATTAATTGATTTGCCATTATAGCTATATCAATAAATAATTCAGCT 751
Db      2374  ATGCTGCGCTGCTCCGTGATCTGGATCCGATTATGTTGCGATATAAAATATCATCTGGCT 2433
QY      752  AAGGAAACATCTTTCAAAGCCCTTATAGTTGCTGCACAGCCAGCTTCATACCTGGGGTCCA 811
Db      2434  CCTGGAGATTTATTTAGAGAGGCTGAAGTTGATAGTCTGCCACAGCCGAGACTGGGGCCCA 2493
QY      812  TACTGGAACAACATCTGGGGGAAAAAGATATATAAGCATGTAGTATTC 858
Db      2494  TTCTTAGCTACAGCACCGCGGGGAAACGCTAACAAATATATGATCGACC 2540

RESULT 15
US-09-191-468-89
; Sequence 89, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Callagher, Michael J.
; APPLICANT: Burgess, Lloyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 123110S01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(450)
; OTHER INFORMATION: Seq 42(E)S2 - HsPC-2; nt 1942-2394; nt 1949 is A
; OTHER INFORMATION: not T (consensus); nt 1959 is C not T (consensus);
; US-09-191-468-89

```

Query Match	9.38;	Score 101.4;	DB 4;	Length 453;
Best Local Similarity	55.78;	Pred. No. 7.1e-18;		
Matches 220;	Conservative 0;	Mismatches 166;	Indels 9;	Gaps
QY	464	GAGCTGTTGGAATCATCTGATTTATATGGAGGAAACAGATTTCATATGAGATACAGAAATG	52	
Db	1	GACCTGAGAGGGATCTCCATTATGTGATATGCTTGCAAAAGATTGTGGAAATTTAAAGATG	60	
QY	524	ATGATGGAGCAAGAGGTGATTAATTCGGCTATGATGATGAGACCTTGCTGGTTTGAAT	58	
Db	61	ATGATTGGATTCACACCTAACATCTTC-----TGAAGTGTGCTGGGCAATTGTA	11	
QY	584	ACGCTATCTCTTTGATTTGCAATATTATACCTGATATGGTGCATTTTCATAGACCTAAT	64	
Db	112	ACCCACACATTTTAAACCTTATCTCTTGCTTCACCTTTTAAACATGGGAACCCATGACC	17	
QY	644	TATGGCGCAATTCATACCCCTGACTGAGGAGATGGCTTATGAGCTGGTGTATGATTGTTTC	70	
Db	172	TATGGCTCTTACCGGTACCCTAACGTCACATGATGGTGTGGATGGCTAATCTGCGCTGT	23	
QY	704	TGCATTATTTGGATTCCAAATATATGGCTATTCATATAAAATATTCAGGCTAAAGAAATC	76	
Db	232	TCCGTCATCTGATGCCATTCATATGTTGTGTATATAAAATGCATTCGGGCCCTGGAAAGATT	291	
QY	764	TTTCACGCCCTTATATAGTTGCTGCAGACACACTCTTAAGTGGGTCATACCTGAGAACAA	823	
Db	292	ATTGAGAGGCTGAAGTTGGTGTGCTGCGCACACGCCGACACTGGGGCCCATTCCTTAAGCTCAA	351	

Wed Aug 27 18:31:19 2003

us-09-940-227-15.rni

Page 10

```

QY      824 CATCGTGGGAAAAGATATTAACACATGTAAGTCC 858
        || || || || || || || || || || || ||
Db      352 CACCGGGGGAGCGTTACAGACATGATCGACC 386

```

Search completed: August 27, 2003, 12:08:56
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 11:20:56 ; Search time 330 Seconds
(without alignments)
7605.789 Million cell updates/sec

Title: US-09-940-227-15

Perfect score: 1094
Sequence: 1 gttcacaggagactgtacc.....tggttttttaaaaaaaa 1094

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCMS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	28.1	1232	13	US-10-027-632-101128
2	307	28.1	1232	13	US-10-027-632-101128
3	147.8	13.5	2863	9	US-09-795-232-1
4	65.4	6.0	553	10	US-09-878-178-1740
5	65.4	6.0	553	13	US-10-046-935-1740
6	65.4	6.0	553	14	US-10-146-502-1740
7	63.2	5.8	1800	12	US-10-345-680-6
8	63.2	5.8	2288	12	US-10-345-680-4
9	56.2	5.1	1911	12	US-10-289-161A-1
10	56.2	5.1	1911	12	US-10-289-161A-2
11	53.4	4.9	1938	9	US-09-845-908-3
12	50.4	4.6	2190	12	US-09-795-693-6
13	50.4	4.6	2190	14	US-10-156-239-6
14	50.4	4.6	2190	14	US-10-199-485-6
15	50.4	4.6	2204	9	US-09-741-149-1
16	50.4	4.6	2304	12	US-10-385-614-1

17	50.4	4.6	3103	9	US-09-795-693-4	Sequence 4, Appl1
18	50.4	4.6	3103	14	US-10-156-239-4	Sequence 4, Appl1
19	50.4	4.6	3103	14	US-10-199-485-4	Sequence 4, Appl1
20	50.2	4.6	424	10	US-09-960-352-7432	Sequence 7432, Ap
21	48.8	4.5	1991	12	US-09-845-908-9	Sequence 9, Appl1
22	47.8	4.4	2558	12	US-10-205-219-31	Sequence 31, Appl1
23	47.6	4.4	1833	10	US-09-861-846-1	Sequence 1, Appl1
24	45.4	4.1	3945	12	US-10-101-510-551	Sequence 551, App
25	45.4	4.1	3946	12	US-10-241-220-53	Sequence 53, App
26	45.4	4.1	3946	14	US-10-213-948-4	Sequence 5, Appl1
27	45	4.1	2093	12	US-09-845-908-5	Sequence 34, Appl1
28	45	4.1	3972	12	US-10-205-194-34	Sequence 950, App
29	44.6	4.1	18512	12	US-10-311-455-950	Sequence 1, Appl1
30	44.2	4.0	2860	12	US-10-298-974-1	Sequence 1, Appl1
31	44.2	4.0	2860	12	US-09-991-458-1	Sequence 1, Appl1
32	43.6	4.0	1051	12	US-09-845-908-7	Sequence 7, Appl1
33	43.6	4.0	2028	10	US-09-917-800A-1429	Sequence 1429, Ap
34	43.6	4.0	2028	12	US-09-845-908-1	Sequence 9, Appl1
35	43.6	4.0	4956	10	US-09-815-923-9	Sequence 444, App
36	43.2	3.9	6534	12	US-10-311-455-444	Sequence 487, App
37	42.8	3.9	6464	12	US-10-311-455-487	Sequence 1285, Ap
38	42.6	3.9	5858	12	US-10-311-455-1285	Sequence 1121, Ap
39	42.4	3.9	2485	9	US-09-923-444-1	Sequence 1247, Ap
40	42.2	3.9	5452	12	US-10-311-455-1121	Sequence 14, Appl1
41	42.2	3.9	13038	12	US-10-311-455-1247	Sequence 404, App
42	42	3.8	6113	12	US-10-204-708-14	Sequence 646, App
43	42	3.8	6113	12	US-10-311-455-404	Sequence 164, App
44	42	3.8	16287	12	US-10-311-455-646	
45	41.8	3.8	17137	12	US-10-311-455-164	

ALIGNMENTS

RESULT 1
US-10-027-632-101128
Sequence 101128, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 101128
LENGTH: 1232
TYPE: DNA
ORGANISM: Human
US-10-027-632-101128

Query Match 28.1%; Score 307; DB 13; Length 1232;
Best Local Similarity 96.9%; Pred. No. 2.1e-67;
Matches 313; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 764 TTTCACGGCTTAATAGTTGTCGACACCGCTTACTACTGGGTCCTACTGACAA 823
DB 432 TTTCAGCGCTTAATAGTTGTCGACACCGCTTACTACTGGGTCCTACTGACAA 491

QY	824	CATCGTGGGGAAGAAATATATAAAGACATGGTGTCTCTAAAAAGAGCGTGGCCATGAATA	883
QY	824		
Db	492	CATGTGGGGGAAGAAATATAAGACATGGTACATCTTAAAAAGAGCGTGACCATGAATA	551
QY	884	CCTACTGTAGTGGCAGCAGAAAAACCGAATGAGATCTCATTTGAAAAAAATATATGATTG	943
Db	552	CCCTCTGTGTGGCAGCAGAAAAACCGAATGAGATCTCATTTGAAAAAAATATATGATTG	611
QY	944	TATAATGATTTTTTTTGTAGAAATAGGGGGACCCCTTATTTATTTGTGTTAACGAATAG	10030
Db	612	TATAATGATTTTTTTTGTAGAAATAGGGGGAACCTTATTTATTTGTGTTAACGAATAG	671
QY	1004	GAANAATGTACTATATGTCATGATAGGATGATTTTTTTTCCATTTAAGCAGGAATGCA	1063
Db	672	GAANAATGTACTATATGTCATGATAGTGTGATTTTTTTTCAATTAAAGCAGGAATGCA	731
QY	1064	ATATAAANAATGTGGTTTTTTTAA	1086
Db	732	ATATAAANAATGTGAATCTCTTAA	754

	RESULT 2	
DB	US-10-027-632-101129	
	; Sequence 101129, Application US/10027632	
	; GENERAL INFORMATION:	
	; APPLICANT: Wang, David G.	
	; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide	
	; TITLE OF INVENTION: Polymorphisms in the Human Genome	
	; FILE REFERENCE: 108827.129	
	; CURRENT APPLICATION NUMBER: US/10/027.632	
	; CURRENT FILING DATE: 2002-04-30	
	; PRIOR APPLICATION NUMBER: US 60/218,006	
	; PRIOR FILING DATE: 2000-07-12	
	; PRIOR APPLICATION NUMBER: US 60/198,676	
	; PRIOR FILING DATE: 2000-04-20	
	; PRIOR APPLICATION NUMBER: US 60/193,483	
	; PRIOR FILING DATE: 2000-03-29	
	; PRIOR APPLICATION NUMBER: US 60/185,218	
	; PRIOR FILING DATE: 2000-02-24	
	; PRIOR APPLICATION NUMBER: US 60/167,363	
	; PRIOR FILING DATE: 1999-11-23	
	; PRIOR APPLICATION NUMBER: US 60/156,358	
	; PRIOR FILING DATE: 1999-09-28	
	; PRIOR APPLICATION NUMBER: US 60/146,002	
	; PRIOR FILING DATE: 1999-08-09	
	; NUMBER OF SEQ ID NOS: 325720	
	; SOFTWARE: FastSeq for Windows Version 4.0	
	; SEQ ID NO 101129	
	; LENGTH: 1232	
	; TYPE: DNA	
	; ORGANISM: Human	
	US-10-027-632-101129	
	Query Match	28.1%; Score 307; DB 13; Length 1232;
	Best Local Similarity	96.9%; Pred. No. 2.1e-67;
	Matches 313; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
OY	764 TTTCAGCCCTTATAGTTGCTGCAGACACCTTACTGGGGCCATTACCTGGAAACA	823
DB		
	432 TTTTCGCCCTTATATAGTTGCTGCAGACACCTTTACTGGGGCTCACTCGAACA	491
OY	824 CATCGTGGAAGAAGATATAAGACATGTGTACTTCTATAAAAAGAGGCTGGCCATGAATA	883
DB		
	492 CATCGTGGAAGAAGATATAAGACATGTGTACTTCTATAAAAAGAGGCTGGCCATGAATA	551
OY	884 CCTACTGTTAGTGGCAGCAGAAAAACCGGAATGAGATCTCATTTGAATAAAATATATGATTG	943
DB		
	552 CCTACTGTTAGTGGCAGCAGAAAAACCGGAATGAGATCTCATTTGAATAAAATATATGATTG	611
OY	944 TATATATGATTTTTTTTAGAATAGGGGACCCTTATTTTGTGTTAACGTAATG	1003
DB		
	612 TATATATGATTTTTTTTAGAATAGGGGAACTTATTTTGTGTTAACGTAATG	671

QY	1004	GAAGATGACACTACTATGTCATCATGATGAGGGATTTTTCCTCCATTTTAAGCAGGAAATGCA	1063
QY	1004	GAAGATGACACTACTATGTCATCATGATGAGGGATTTTTCCTCCATTTTAAGCAGGAAATGCA	1063
Db	672	GAAGATGACACTACTATGTCATCATGATGAGGGATTTTTCCTCCATTTTAAGCAGGAAATGCA	731
QY	1064	ATATATATATATGCGTTTATTTTAA	1086
QY	1064	ATATATATATGCGTTTATTTTAA	1086
Db	732	ATATATATATGCGATTCCTTAA	754

```

RESULT 3
US-09-795-232-1
; Sequence 1, Application US/09795232
; Patent No. US20010012627A1
; GENERAL INFORMATION:
; APPLICANT: Anthony M. Brown
; APPLICANT: Conrad Gerald Chapman
; APPLICANT: Israel Simon Gloger
; APPLICANT: Joanne Rachel Evans
; APPLICANT: William Cairns
; APPLICANT: Hugh Jonathan Herdon
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30176-D1
; CURRENT APPLICATION NUMBER: US/09/795,232
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/182,728
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 9818890.7
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2863
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-795-232-1

Query Match      13.5%; Score 147.8; DB 9; Length 2863;
Best Local Similarity 54.7%; Pred. No. 6.3e-27;
Matches 321; Conservative 0; Mismatches 257; Indels 9; Gaps 1;

272  GAAACGATACACACACAAACATTCAGATTATTTCCCAAAAGTATGAAGAAAATAGGGTT 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2005  GAGACCAATAGAGACCTCATCATCGACAGACGAGTTCCCAAGTACCTACGACACACAGGCA 2064
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
332  CCCATATCTTTGGGCGTCTGCTGTTTGTTCCTTCCTTGGTCMCGTGTGTCAGTCAG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2065  GTGTTTACTCTGGGCTGCTGTCGATTGTTTCTTCATCAGGGTTTCCATATGATCACTCAG 2124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
392  GCTGGAATTTACTGGGTCATCTGATGTGACCACTTCTGCTGATGAGGGGCATTTTAAT 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2125  GGTGAATTTACATGTTTCACGTCGTGTGACACGCTATGCTGCTCTATGACCTTGTCATC 2184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
452  GCAGCTATCTGAGGCTAGTTGGAATCATCTGGAATTATGAGGGGACAGATTCATGTAG 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2185  ATTGCATTTTGAAGCTCGTGGGGAATCTCTATGTAATGAGCTTGCAAAAGATTCGTGAA 2244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
512  GATACAGAAATCATGATGATGAGCAAGAAAGGAGATTCGAGCTATGAGTGAGAGACTGCG 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2245  GATATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
572  TGGTTTGTAATATACGCTATTCCTTTTGATGCAATATTTATTCGTGTCATGAGTGCATTT 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2256  TGGCGATTTGTAACCCCAACCATTTTAACCTTTATCTGTTGCTTCAGCTTTTACCAATG 2355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
632  CATAGACCTAATTATGCGGCAATTCATACCTGACTGGGAGTGGCTTGAAGCGTGCT 691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2356  GAGCCCATAGACCTAATGAGGCTCTTACGCTATACCTAATGCTCAGTGCAGTGCAGTGC 2415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
692  ATGATGTTTTCGTGATATTTGATGATTCATATTTATGCTATCATCAAAATATTCAGGCT 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2416  ATGCTGCGCTGTTCCGTCATCTGGAATCCCAATATGTTTGATGATAAATGATCATCTGGCC 2475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


QY 752 AAGGAACATCTTTCAACGCCCTTAAGTTGCTGACAGACGCTTAACTGGGGTCCA 811
Db 2476 CCTGAGAGATTATTGAGAGAGCTGAGTGGCTGCGCCAGCGGAGTGGGGCCCA 2535
QY 812 TACCTGGAACACATCTCGTGGGGAAGAATATAAACAATGAGTCC 858
Db 2536 TTCTTACCTCAACACGCGGGGAGCGTTACAAAGAATGATGACGCC 2582

RESULT 4
US-09-878-178-1740
: Sequence 1740, Application US/09878178
: Patent No. US20020177552A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Secretist, Heather
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
: FILE REFERENCE: 210121.527
: CURRENT APPLICATION NUMBER: US/09/878.178
: CURRENT FILING DATE: 2001-06-08
: NUMBER OF SEQ ID NOS: 2237
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1740
: LENGTH: 553
: TYPE: DNA
: ORGANISM: Homo sapien
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(553)
: OTHER INFORMATION: n = A,T,C or G
US-09-878-178-1740

Query Match
Best Local Similarity 92.0%; Score 65.4; DB 10; Length 553;
Matches 69; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1012 ACATACATATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGAAATGCAATATATAA 1071
Db 1 ACATACATATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGAAATGCAATATATAA 60

QY 1072 ATGTGCTTTTAA 1086
Db 61 ATGTGAATCTCTTAA 75

RESULT 5
US-10-046-935-1740
: Sequence 1740, Application US/10046935
: Publication No. US20020156011A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Secretist, Heather
: APPLICANT: Wang, Aijun
: APPLICANT: Stolk, John A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
: FILE REFERENCE: 210121.527C1
: CURRENT APPLICATION NUMBER: US/10/046.935
: CURRENT FILING DATE: 2002-01-15
: NUMBER OF SEQ ID NOS: 2239
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1740
: LENGTH: 553
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 546
: OTHER INFORMATION: n = A,T,C or G
US-10-046-935-1740

QY 1012 ACATACATATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGAAATGCAATATATAA 1071
Db 1 ACATACATATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGAAATGCAATATATAA 60

QY 1072 ATGTGCTTTTAA 1086
Db 61 ATGTGAATCTCTTAA 75

RESULT 6
US-10-146-502-1740
: Sequence 1740, Application US/10146502
: Publication No. US20030069180A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Secretist, Heather
: APPLICANT: Wang, Aijun
: APPLICANT: Stolk, John A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
: FILE REFERENCE: 210121.527C2
: CURRENT APPLICATION NUMBER: US/10/146.502
: CURRENT FILING DATE: 2002-05-14
: NUMBER OF SEQ ID NOS: 2241
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1740
: LENGTH: 553
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 546
: OTHER INFORMATION: n = A,T,C or G
US-10-146-502-1740

Query Match
Best Local Similarity 92.0%; Score 65.4; DB 14; Length 553;
Matches 69; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1012 ACATACATATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGAAATGCAATATATAA 1071
Db 1 ACATACATATGTTTCATGATAGGATGATTTTCCCATTTAAGCAGAAATGCAATATATAA 60

QY 1072 ATGTGCTTTTAA 1086
Db 61 ATGTGAATCTCTTAA 75

RESULT 7
US-10-345-680-6
: Sequence 6, Application US/10345680
: Publication No. US20030148394A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: APPLICANT: Siles-Santiago, Inmaculada
: APPLICANT: Venkateswarlu, Karicheti
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
: TITLE OF INVENTION: UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,
: TITLE OF INVENTION: 641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,
: TITLE OF INVENTION: 12303, 985, 13237, 13601, 18926, 318, 2058 OR 6351 MOLECULES.
: FILE REFERENCE: MP102-012P1RM.OM1
: CURRENT APPLICATION NUMBER: US/10/345.680
: CURRENT FILING DATE: 2003-01-16
: PRIOR APPLICATION NUMBER: US 60/349,511
: PRIOR FILING DATE: 2002-01-18
: PRIOR APPLICATION NUMBER: US 60/360,500
: PRIOR FILING DATE: 2002-02-28

```

? PRIOR APPLICATION NUMBER: US 60/365,041
? PRIOR FILING DATE: 2002-03-15
? PRIOR APPLICATION NUMBER: US 60/374,063
? PRIOR FILING DATE: 2002-04-19
? PRIOR APPLICATION NUMBER: US 60/403,468
? PRIOR FILING DATE: 2002-08-14
? PRIOR APPLICATION NUMBER: US 60/414,262
? PRIOR FILING DATE: 2002-09-27
? PRIOR APPLICATION NUMBER: US 60/419,986
? PRIOR FILING DATE: 2002-10-21
? PRIOR APPLICATION NUMBER: US 60/423,809
? PRIOR FILING DATE: 2002-11-05
? PRIOR APPLICATION NUMBER: US 60/429,797
? PRIOR FILING DATE: 2002-11-26
? NUMBER OF SEQ ID NOS: 66
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 6
? LENGTH: 1800
? TYPE: DNA
? ORGANISM: Homo Sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)...(1800)
? US-10-345-680-6

```

Query Match	5.8%;	Score 63.2;	DB 12;	Length 1800;
Best Local Similarity	49.4%;	Pred. No. 1.2e-05;		
Matches 229;	Conservative 0;	Mismatches 223;	Indels 12;	Gaps 2;

QY 260 TCCCCCTCGAAGAAACGATCACACAACAATTCCAGATTTATTTCCCAGAAGTGTGAG 319
 | | | | | | |
Db 1192 TTTCGCACCTGTGGAGGCTTCATCATCACCACCCCTGTGTGATGAGTACCCCAAGCTCCTCCG 1251

QY 320 AAAATGAGGTTCCCATACCTTTGGGCTGCTGCTGTTTGTCTCCTTGGTCTGTC 379
 Db 1252 AACCGAGGAGCTCTTCTATTTGCTGTGCTGTGCATCATCTCCATCCTGATCGGCTCTCT 1311

QY 380 TGTGTGACTCAGGCTGGAAATTACTGGGTCATCTGATTGACCACTTCTGTGC---TGA 436

Db 1312 AACATCACCAGGGGGTATTTTAGTCTTCAACCTCTTTGACTACTACTGTGCAGTGGC 1371

QY 437 TGGGGCATTTTAATTGACAGTAACTGAGACTAGTTGGAAATCAATCGAATTTATGAGAGG 496

Db 1372 ATGAGGCTCTGTTCTCTGTTCTTTTGAATGAGTCTCAATTCCTGGTTTACGGTGC 1431

497 AACAGATTCATTGAGGATACAGAAATGATGATTGGAGCAAAAGAGGTGATATTCTGGCTA 556

QY

1432 ACCGATTCATGACCAATATCTCCACAGAGATGGTTGGATCCAGAGGCCCTGCATC----- 1483

Db

[illegible][illegible]

677 GCCTTAGCCGTGATGATGTTCTGCAATFATTGGATCC 720

RESULT 8

```

; Sequence 4, Application US/10345680
; Publication No. US200301148394A1
; GENERAL INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

TITLE OF INVENTION:	641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,
TITLE OF INVENTION:	UROLOGICAL DISORDERS USING 14C, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842,

```

1  TITLE OF INVENTION: 12303, 985, 13237, 13601, 18926, 318, 2058 OR 6351 MOLECULES
2  FILE REFERENCE: MP102-012P1RM.OMNI
3  CURRENT APPLICATION NUMBER: US/10/345, 680
4  CURRENT FILING DATE: 2003-01-16
5  PRIOR APPLICATION NUMBER: US 60/349,511
6  PRIOR FILING DATE: 2002-01-18
7  PRIOR APPLICATION NUMBER: US 60/360,500
8  PRIOR FILING DATE: 2002-02-28
9  PRIOR APPLICATION NUMBER: US 60/365,041
10 PRIOR FILING DATE: 2002-03-15
11 PRIOR APPLICATION NUMBER: US 60/374,063
12 PRIOR FILING DATE: 2002-04-19
13 PRIOR APPLICATION NUMBER: US 60/403,468
14 PRIOR FILING DATE: 2002-08-14
15 PRIOR APPLICATION NUMBER: US 60/414,262
16 PRIOR FILING DATE: 2002-09-27
17 PRIOR APPLICATION NUMBER: US 60/419,986
18 PRIOR FILING DATE: 2002-10-21
19 PRIOR APPLICATION NUMBER: US 60/423,809
20 PRIOR FILING DATE: 2002-11-05
21 PRIOR APPLICATION NUMBER: US 60/429,797
22 PRIOR FILING DATE: 2002-11-26
23 NUMBER OF SEQ ID NOS: 66
24 SOFTWARE: FastSeq for Windows Version 4.0
25 SEQ ID NO 4
26 LENGTH: 2298
27 TYPE: DNA
28 ORGANISM: Homo Sapiens
29 FEATURE:
30 NAME/KEY: CDS
31 LOCATION: (235)...(2034)
32 US-10-345-680-4

```

Query Match	5 8%	Score 63.2:	DB 12:	Length 2298:
Best Local Similarity	49.14%	Pred. NO. 1.4e-05:		
Matches 229:	0:	Mismatches 223:	Indels 12:	Gaps 2:

QY	260	1000	CTCTAAGAAACGATCAACACAAATCAATGATTTATTTCCCAAGATGTAAG	319
Db	1426	TTCTGCACTGTGGAGGGCTTCATCAACACGCCCTGGTGGATAGTAACCCAGGCTCTCCGC	1485	

QY 320 AAAATGAGGGTTCCTCAATACCTTTGGGCTGCTGCTTGGTTTGTTCCTCTGCTGCTC 379
 1486 AACCGAGAGAGCTTCTTCATTTGGCTGCTGCTGCATATCTCTACCTGATGGGCTCTCT 1545

Qy 380 TGTGTACTCAGCGCTGGAATTTCCTGGGTTCATCTGATTCACCACTTCTGTGC--TGGC 436
 1546 AACATCACATGAGGGGGTATTTCATCTTCAACCTCTTTGCTACTTACCTCTGCAGATGGC 1605

QY 437 TGGGCGATTTTAAATTGCAGCTAATACGTGAGCTAGTTGGAAATCATTCGTGAATTTATGAGCG 496
| | | | | | | | | | | | | | | |
Db 1606 ATGACGCTGCTGCTCCTCCCTGCTCTTGCAATGCGTCCTPPTTCCTGAGTTTACGGTGC 1665

Oy AACAGATTCATTGAGATACAGAAATGTGATTGGACAAGAGTGATATTCTGCCTA 556
| | | | |
Oy 497 AACCAGTTCATTGAGATACAGAAATGTGATTGGACAAGAGTGATATTCTGCCTA 556
| | | | |
Db ACCCATTTCAATGCCAATTTCCGAAGAAGATGGTGGATCCTCGGCCCTGCATC----- 1716
| | | | |

[illegible]

617 TCATTGGTCATTTGATGACCTAATTATGCGCATTCGATACCGTGAGGAGTT 676

QY 677 GCTTAGCCGTCGTATCATTTTTCGCATTATTGGATTC 720
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |

RESULT 9

US-10-205-101A
; Sequence 1, Application US/10289161A

```

; Publication No. US20030152970A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods and Compositions to Treat Pain
; TITLE OF INVENTION: and Painful Disorders Using 577, 20739 or 57145
; FILE REFERENCE: MPI2001-287PIR(M)
; CURRENT APPLICATION NUMBER: US/10/289,161A
; PRIOR FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/333,073
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-289-161A-1

```

```

Query Match      5.18; Score 56.2; DB 12; Length 1911;
Best Local Similarity 44.9%; Pred. No. 0.00076;
Matches 266; Conservative 0; Mismatches 318; Indels 9; Gaps 1;

```

```

OY 264 CCTCTGAGAAAGATCAACAACATTCAGATTATTTCCCAAGTATGAAGAAA 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1205 CCTTCTGAGAGACCATTTGTGACAGCTGTGACAGATGAGTTCCCTACTACCTGGGCCCA 1264
OY 324 TGAGGGTTCCCATTAACCTTTGGGCTGCTGCTTTGTTTCTCCTGCTGCTGCTG 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1265 AGAAGGGGCTGTTCTCAGGGCTCATCTGCGGCGCATCTACCTGATGCTGCTCA 1324
OY 384 TGACTCAGGCTGGAATTTACTGGGTTTATCTGATTCACACTTCTGCTGATGGGCA 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1325 CCACTATGAGGGGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1384
OY 444 TTTTAATTGAGCTATACATGAGCTAGTGGAAATCACTGATTTATGAGGAAACAT 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 TGATGCTGCTGCTTATACCAACATGCTTCCCTGACAGGGGTATGCAATTCAGAGGT 1444
OY 504 TCATTGAGATACAGAAATGATGAGCAAAAGAGTGAATTTCTGCTATGGTGA 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1445 TCTGCCAGACATCCACATGATGCTGGGCTTCAAGCGGGGCTCTACTTC-----A 1495
OY 564 GAGCTGCTGCTTTGTAATTAACGCTATCTTTGATTCGAATTTATCTGCTATGG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1496 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1555
OY 624 TGCATTTTATAGACCTAATTTATGCGCAATTCATACCTGAGGGGCTGCTTTAG 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1556 TCAGATACCAAGCCCTGCGAGTATGCGATTCACGCTTCCGCGCTGGGCTGAGCTGCTGG 1615
OY 684 GCTGCTGATGATTTGTTCTGCAATTTTGGATTCGAATTTATGCTATCAATAAATA 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1616 GCATCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1675
OY 744 TTCAGGCTAAAGAAACATCTTTCAAGCGCTTTAAGTTGCTGCAACACGCTTCTACT 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1676 TTCGAAAGAGGGCTGCTGCTGAGGCGCTCCAAAGGCGACCGCGGCGCATGAGACT 1735
OY 804 GGGGTCATACCTGGAACAACATCGTGGGAAAGATATAAGCATGAGTGT 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1736 GGGGACCATCTGCTGAGAGAAACCGAGCGGCTATGCTGCGCAGCTGGCT 1788

```

```

RESULT 10
US-10-289-161A-2
; Sequence 2, Application US/10289161A
; Publication No. US20030152970A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods and Compositions to Treat Pain
; TITLE OF INVENTION: and Painful Disorders Using 577, 20739 or 57145

```

```

; FILE REFERENCE: MPI2001-287PIR(M)
; CURRENT APPLICATION NUMBER: US/10/289,161A
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/333,073
; PRIOR FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1911)
US-10-289-161A-2

```

```

Query Match      5.18; Score 56.2; DB 12; Length 1911;
Best Local Similarity 44.9%; Pred. No. 0.00076;
Matches 266; Conservative 0; Mismatches 318; Indels 9; Gaps 1;

```

```

OY 264 CCTCTGAGAAAGATCAACAACATTCAGATTATTTCCCAAGTATGAAGAAA 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1205 CCTTCTGAGAGACCATTTGTGACAGCTGTGACAGATGAGTTCCCTACTACCTGGGCCCA 1264
OY 324 TGAGGGTTCCCATTAACCTTTGGGCTGCTGCTTTGTTTCTCCTGCTGCTGCTG 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1265 AGAAGGGGCTGTTCTCAGGGCTCATCTGCGGCGCATCTACCTGATGCTGCTCA 1324
OY 384 TGACTCAGGCTGGAATTTACTGGGTTTATCTGATTCACACTTCTGCTGATGGGCA 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1325 CCACTATGAGGGGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1384
OY 444 TTTTAATTGAGCTATACATGAGCTAGTGGAAATCACTGATTTATGAGGAAACAT 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 TGATGCTGCTGCTTATACCAACATGCTTCCCTGACAGGGGTATGCAATTCAGAGGT 1444
OY 504 TCATTGAGATACAGAAATGATGAGCAAAAGAGTGAATTTCTGCTATGGTGA 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1445 TCTGCCAGACATCCACATGATGCTGGGCTTCAAGCGGGGCTCTACTTC-----A 1495
OY 564 GAGCTGCTGCTTTGTAATTAACGCTATCTTTGATTCGAATTTATCTGCTATGG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1496 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1555
OY 624 TGCATTTTATAGACCTAATTTATGCGCAATTCATACCTGAGGGGCTGCTTTAG 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1556 TCAGATACCAAGCCCTGCGAGTATGCGATTCACGCTTCCGCGCTGGGCTGAGCTGCTGG 1615
OY 684 GCTGCTGATGATTTGTTCTGCAATTTTGGATTCGAATTTATGCTATCAATAAATA 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1616 GCATCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1675
OY 744 TTCAGGCTAAAGAAACATCTTTCAAGCGCTTTAAGTTGCTGCAACACGCTTCTACT 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1676 TTCGAAAGAGGGCTGCTGCTGAGGCGCTCCAAAGGCGACCGCGGCGCATGAGACT 1735
OY 804 GGGGTCATACCTGGAACAACATCGTGGGAAAGATATAAGCATGAGTGT 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1736 GGGGACCATCTGCTGAGAGAAACCGAGCGGCTATGCTGCGCAGCTGGCT 1788

```

```

RESULT 11
US-09-845-908-3
; Sequence 3, Application US/09845908
; Publication No. US20030143729A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Kell E. et al
; TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 40558-D
; CURRENT APPLICATION NUMBER: US/09/845,908
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 28

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 09:55:34 ; Search time 2773 Seconds
(without alignments) 9588.567 Million cell updates/sec

Title: US-09-940-227-15

Perfect score: 1094
Sequence: 1 gttcacaggaggactgtacc.....tggtttttaaaaaaaaaa 1094

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estb9:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrl:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	621.8	56.8	665	13	Bu687102 UI-CF-DU1
C 2	611.8	55.9	695	9	AM190954 x143b12.x
C 3	561	51.3	2199	11	AK018553 Mus muscu
C 4	532.8	48.7	574	9	AI669617 tw34e04.x

C 5	528	48.3	617	9	AM812970
C 6	507	46.3	530	12	BM826722 K-EST009
C 7	476.6	43.6	691	12	BG964290 602829090
C 8	472.6	43.2	487	9	AA552658
C 9	468.8	42.9	501	10	BF096011
C 10	425.2	38.9	614	14	CB579897
C 11	415	37.9	567	4	BX523178
C 12	411.6	37.6	582	14	CB586665
C 13	381	34.8	484	10	BF774564
C 14	344.2	31.5	438	9	AI429024
C 15	325.8	29.8	360	10	BE003214
C 16	324	29.6	460	13	BQ378904
C 17	290	26.5	332	10	BE002150
C 18	276	25.2	463	9	AI605513
C 19	273.8	25.0	386	9	AA592728
C 20	267.6	24.5	682	9	AL871486
C 21	246.6	22.5	636	9	AL893341
C 22	222	20.3	893	13	BH899741
C 23	219.4	20.1	889	14	CA791412
C 24	218	19.9	374	10	BE819593
C 25	213	19.5	337	13	BQ082831
C 26	209	19.1	498	9	AL872634
C 27	207.4	19.0	482	13	BQ311137
C 28	205.6	18.8	282	10	BE183155
C 29	205.6	16.1	201	13	BQ378913
C 30	176.4	15.1	1068	12	BM559743
C 31	125.4	11.5	1068	13	BM302794
C 32	124	11.0	740	14	CA312866
C 33	120	9.5	142	10	BE005236
C 34	104.4	8.1	264	9	AV372892
C 35	101.8	8.1	574	9	AM645092
C 36	88.4	8.0	553	10	BG467430
C 37	87.4	7.6	345	13	BX095910
C 38	83	7.5	151	10	BE814811
C 39	82	7.3	660	13	BM045601
C 40	80.2	7.2	806	13	BM017451
C 41	78.8	7.0	641	13	BU140319
C 42	77	7.0	787	12	BG781704
C 43	76.4	6.8	757	13	BU141486
C 44	74.2	6.6	234	10	BU141317
C 45	72.2				

ALIGNMENTS

RESULT 1
BU687102/c 665 bp mRNA linear EST 07-OCT-2002
UI-CF-DU1-ado-b-18-0-UI-s1 UI-CF-DU1 Homo sapiens CDNA clone
DEFINITION
UI-CF-DU1-ado-b-18-0-UI 3', mRNA sequence.
ACCESSION
BU687102
VERSION
BU687102.1 GI:23542655
KEYWORDS
EST.

SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

JOURNAL
MEDLINE
PUBMED
TITLE
Normalisation and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548

COMMENT
McCrack, McCray, PB

McCrack Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@iowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cdDNA library preparation: Dr. M. Bento Soares, University of Iowa
cdNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).

Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES	Location/Qualifiers
1	655

Source

```

source
1. 665
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon.9606"
  /clone="UI-CF-DU1-ado-b-18-0-UI"
  /tissue_type="Primary Lung Epithelial Cells"
  /dev_stage="Adult"
  /lab_host="DH10B (Life Technologies) (TI phase resistant)"
  /clone_lib="UI-CF-DU1"
  /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
  modified polylinker; Site.1: EcoR I; Site.2: Not I;
  UI-CF-DU1 is a normalized cDNA library containing the
  following tissue(s): Primary Lung Epithelial Cells The
  library was constructed according to Bonaldi, Lennon and
  Soares, Genome Research, 6:791-806, 1996. First strand
  cDNA synthesis was primed with an oligo-dT primer
  containing a Not I site. Double stranded cDNA was ligated
  to an EcoR I adaptor, digested with Not I, and cloned
  directionally into pT73-Pac vector. The oligonucleotide
  used to prime the synthesis of first-strand cDNA contains
  a library tag sequence that is located between the Not I
  site and the (dT)18 tail. The sequence tag for this
  library is GGCTGTAGGC.
  TAG_Lib-UI-CF-DU1
  TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
  TAG_SEQ=GGCTGTAGGC"
209 a 148 c 92 g 216 t

```

[illegible]

QY 920 CTCATGTGAAAAAATATATGATTTGTATATATGTGATTTTTTTTAAATAGGGGGACCCCTTA 979
Db 183 CTCATTTGAAAAAATAATATGATTTGTATATATGTGATTTTTTTTAAATAGGGGGACCTTA 124			
QY 980 TTTATTTGTGTGTTAACTGAATAGAGAAATGTACATCTACTATGTCATGATAGGCTGATTT 1039			
Db 123 TTTATTTGTGTGTTAACTGAATAGAGAAATGTACATCTACTATGTCATGATAGGCTGATTT 64			
QY 1040 TTTATTTGTGTTAACTGAATAGAGAAATGTGCTTTTTTTTTTAAAAA 1094			
Db 63 TTTTCACCTTTTAAAGCAGGAATGCATATATAAAAAATGTGAATCTCTTAAAAA 9			

RESULT 2				
AW190954/c				
LOCUS	AM190954	695 bp	mRNA	linear EST 22-NOV-1999
DEFINITION	x13b12.x1 NCT-CGAP Panl Homo sapiens cDNA clone IMAGE:2677439 3'			
	similar to TR:O15003 O15003 GILT-1 LIKE ; ,			mRNA sequence.

ACCESSION	AW190954	GI:6465434
VERSION	AW190954.1	

KEYWORDS	EST.
SOURCE	Homo sapiens (human)

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Craniiata; Mammalia; Primates; Hominidae; Homo
----------	---

REFERENCE
Mammalia; Eutheria; Primates; Caldiiniini; hominidae; Homo.
1 (bases 1 to 695)

AUTHORS NCI-CGAP <http://www.nci.nih.nlm.nih.gov/hot/cgap/>
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
URL <http://www.nci.nih.gov/ncicg>

Tumor Gene Index
Unpublished
Journal
Copyright
Contact: Robert Strausberg, Ph.D.

COMMENT:
Contact: Robert Straussberg, rns@nslc.gov
Email: cgabbs-r@mail.nih.gov
Life Technologies Catalog #: 1154

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCRP clone distribution information can be found at: <http://ccr.cancer.gov>

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/pbrr/image/image.html

Seq primer: -40UP from Glibco
High quality sequence stop: 419.

FEATURES	Location/Qualifiers
source	1-695

```

/organism="Homo sapiens"
/mol type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:2677439"

```

```

/tissue_type="adenocarcinoma"
/lab_host="DH10B"

```

```

/clone_lib="NCI_CGAP_pan1"
/organ="pancreas; Vector: pCMV-SPORT6; site_1: SalI

```

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:

	11548-013"	
BASE COUNT	224 a	154 c
		105 g
		212 t

ORIGIN

Query Match	Score	DB	Length
Best Local Similarity	94.48;	Pred. No. 1.1e-120;	

Matches	656;	Conservative	0;	Mismatches	37;	Indels	2;	Gaps	2
---------	------	--------------	----	------------	-----	--------	----	------	---

401 TACTGGCTTCATGCTGATTGACCACTTCGTGCTGATGGGCAATTAAATTCACCTATA 460

Db 695 TACTGGGTTTCATCTGATTTGA-CAC TTCTGTAGCTGGATGGGCATTTTATTTGACAGGCTTAC 630

461 CTGGAGCTTGTGGAATCATCTGGATTATGAGGAGACAGATTCTATGGAGGATACAGAA 520

Db 636 CTGAGCGCTAGTGGAAATCATTCIGATTATTGAGGAGCACCAGATTCATTGAGGATACAGAA 5/

521 ATGATGATTGGAGCAAGAGGTCATATCTGGCATGTGGAGAGCGTGGCTGGATTGTA 580

Db 576 ATGATGATTGGACAAAGAGGIGGATATTCCTGGCCTAGTGGTGGAGAGCGCTTCTGGTTTGTA 317

581 ATTACGCCATCTTTGATTCGACATATTATCTGTGATTTGGTGCAGTTTCATGACCT 640


```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:9030613J17"
/db_xref="MGI:1907717"
/db_xref="taxon:10090"
/clone="9030613J17"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
108..2024
/note="unnamed protein product; NA+ AND CL-COUPLED NEUTRAL
AND BASIC AMINO ACID TRANSPORTER AIBO,+ (SPR109160,
evidence: FASTV, 99.7%ID, 100%length, match=1914)
putative"
/codon_start=1
/protein_id="BAB31272.1"
/db_xref="GI:12858314"
/db_xref="MGI:1890216"
/translation="MDRLCPNFKRQKROKRYTASSENHVEGENDENRGNWKSXD
VILSWGVAVGLGNWRPELYFTNGGAFILIPVIMLALGLPFLFECISLGFASL
GPRVWRILPLFOGVTIMVLSVFAIYVNIYLSYLFASQSVLPWANGSMA
DENCSNPLVTCGNVSTGNGEPMNISMVNTNCLNGSEVEREPGLPSROYMDKT
LORSSGMDTGVVWVIALCLLAMIIVGAALFKIKSGKRVFTALPPVILILL
IRGATLEGASGISYVIGASNFTKIREAVKDAIOTFISLVSAMGGLVASYNK
FNNGCSDAIIVCLNCLTSVPAFISILGHMAHISKEVSPVAKSFDIATAP
BALAOPAPPEWSILEPFMLITGLDPSASTETTTTFQDLFPAMKRMVPTLGC
CITFLILGLICVQAGIYVWHLIDHAGMGJLIALEIAGIIVIGYGNRPEDIEI
MIGAKRWITFMWMAKCFVITPILSALVMSVYFHRDVEDIYPDWGVALGCMCI
ICILNIPIMAIKTIQVAGNLIQRIISCIPIASNMGPYLEKHREBRYDMAPEAKET
DHEIPTISSTPE"
polyA_signal
2181..2186
/note="putative"
polyA_site
2199
/note="putative"
BASE COUNT 590 a 404 c 507 g 698 t
ORIGIN
Query Match 51.3%; Score 561; DB 11; Length 2199;
Best Local Similarity 81.6%; Pred. No. 9.3e-110;
Matches 675; Conservative 0; Mismatches 145; Indels 7; Gaps 2;

```

```

Db 1792 GCTGCTGATGATCAATTTCTGCATTTATTTGGATACCAATTATGCTATTCATAAATAAG 1851
QY 744 TTCAGGCTAAAGAAACATCTTTCACAGCCTATATAGTGGCGAGACCACTTCTACT 803
Db 1852 TTCAGGCAAGAGAAACATCTTTCACAGCCTATATAGTGGCGAGACCACTTCTACT 1911
QY 804 GGGGTCATACCTCGAACAACATCTGCGGAAAGATATTAAGACATGTAGTCTTAAGA 863
Db 1912 GGGGTCATACCTCGAACAACATCTGCGGAAAGATATTAAGACATGTAGTCTTAAGA 1971
QY 864 AAGAGCGTGGCCATGAATACCTACTGTAGTGGCAGCAGAAACCGGATAGATCTCA 923
Db 1972 AGGAGACTGACCGAAGATACCTACTGTAGTGGCAGCAGAAACCGGATAGATCTCA 2031
QY 924 ---TTGAAAAAATATATGATGTATATATGATATTTTATTAAGTGGGAGACCTTAT 980
Db 2032 CTTTGAATAATATCTAATTTCTTATATATGATTTTGTAGAGTAAAGGATTTTAT 2091
QY 981 TTTATTTGTGTATTAAGTAAAGAAATGTACTACTATGTCATGATAGGT---GA 1036
Db 2092 TTATTTGTATGTTAATCAATTAGGAAATGTACTACTACTACTACTATGTTAG 2151
QY 1037 TTTTTCCTCATTAAGCAGAAATCAATTAATAATGTGTTT 1083
Db 2152 TACTTTTTCATTAATCAAGATGAATATTAATATGATCTAT 2198

RESULT 4
A1669617/c 574 bp mRNA linear EST 15-DEC-1999
LOCUS tw34e04.x1 NCI-CGAP ut1 Homo sapiens cDNA clone IMAGE:2261598.3
DEFINITION similar to TR:015003 015003 GUTR-1 LIKE.; mRNA sequence.
ACCESSION A1669617
VERSION A1669617.1 GI:4834391
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 574)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bbrp/image/image.html
Insert Length: 1364 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 385.
Location/Qualifiers
1..574
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2261598"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI-CGAP-UT1"
/note="Organ: uterus; Vector: pCMV-Sport6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"
BASE COUNT 182 a 118 c 80 g 193 t 1 others
ORIGIN

```

Query Match 48.7%; Score 532.8; DB 9; Length 574;
 Best Local Similarity 99.3%; Pred. No. 9,3e-104;
 Matches 545; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

546 TATTCCTGCTATGCTGAGACGCTGCTGTTGTAATTAAGCCTATCCCTTTGATGCA 605
 573 TATTCCTGCTATGCTGAGACGCTGCTGTTGTAATTAAGCCTATCCCTTTGATGCA 515
 606 TATTCCTGCTATGCTGAGACGCTGCTGTTGTAATTAAGCCTATCCCTTTGATGCA 665
 514 TATTCCTGCTATGCTGAGACGCTGCTGTTGTAATTAAGCCTATCCCTTTGATGCA 455
 666 ACTGGGAGCTGCTGTTAGGCTGCTGTTAGTATGATGTTTTCGATTAATTTGATGCA 725
 454 ACTGGGAGCTGCTGTTAGGCTGCTGTTAGTATGATGTTTTCGATTAATTTGATGCA 395
 726 TGGCTATCATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 785
 394 TGGCTATCATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 335
 786 GCAAGCAGCTTCTTAACCTGGGCTCATACCTGGAGACACATGCTGGGAAAGATTAAG 845
 334 GCAAGCAGCTTCTTAACCTGGGCTCATACCTGGAGACACATGCTGGGAAAGATTAAG 275
 846 ACATGCTAGTCTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 905
 274 ACATGCTAGTCTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 215
 906 AACCGGAATGAGATCTCATTTGAAAAAATATATGATGTAATGATGATTTTGTAGCA 965
 214 AACCGGAATGAGATCTCATTTGAAAAAATATATGATGTAATGATGATTTTGTAGCA 155
 966 TAGGGGAGCCTTATTTATTTGCTGTTACTGTAATGAAATGAAATGATGATGTTCA 1025
 154 TAGGGGAGCCTTATTTATTTGCTGTTACTGTAATGAAATGAAATGATGATGTTCA 95
 1026 TATAGGGGATTTTATTTTCCATTAAGAGAAATGCAATATTAATGATGTTTGTAGCA 1085
 94 TATAGGGGATTTTATTTTCCATTAAGAGAAATGCAATATTAATGATGTTTGTAGCA 35
 1086 AAAAAA 1094
 34 AAAAAA 26

RESULT 5
 AM812970/c 617 bp mRNA linear EST 17-MAY-2000
 LOCUS RC3-ST0186-230300-019-406 ST0186 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION AM812970
 VERSION AM812970.1 GI:7905964
 KEYWORDS EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 DIES Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brlones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
 CONTACT: Simpson A.J.G.
 LABORATORY OF Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=42-RC3-ST0186-230
 300-019-a06413-2000-03-23414-1)
 Seq primer: puc 18 forward
 High quality sequence start: 32
 High quality sequence stop: 617.
 Location/Qualifiers

FEATURES

source
 1. 617
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="ST0186"
 /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 205 a 130 c 95 g 187 t
 ORIGIN

Query Match 48.3%; Score 528; DB 9; Length 617;
 Best Local Similarity 98.0%; Pred. No. 1e-102;
 Matches 545; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

531 GAGAAAGAGTGGATATTTGCTATGCTGATGAGAGCTGCTGTTGTAATTAAGCCTA 590
 617 GAGAAAGAGTGGATATTTGCTATGCTGATGAGAGCTGCTGTTGTAATTAAGCCTA 558
 591 TCCCTTGAATGCAATTTATTCGTCATGCTGCTGATGAGAGCTGCTGTTGTAATTAAGCCTA 650
 557 TCCCTTGAATGCAATTTATTCGTCATGCTGCTGATGAGAGCTGCTGTTGTAATTAAGCCTA 498
 651 CAATTCATACCTGACTGGGAGTCTTATGCTGCTGATGAGAGCTGCTGTTGTAATTAAGCCTA 710
 497 CAATTCATACCTGACTGGGAGTCTTATGCTGCTGATGAGAGCTGCTGTTGTAATTAAGCCTA 438
 711 TTTGATTCATTAATGCTATATATAATAATAATAATAATAATAATAATAATAATAATA 770
 437 TTTGATTCATTAATGCTATATATAATAATAATAATAATAATAATAATAATAATAATA 378
 771 GCCTTAATGCTGCTGACACAGCTTCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830
 377 GCCTTAATGCTGCTGACACAGCTTCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 318
 831 GGGAAAGATTAAGACATGATGTTCTTAATAATAATAATAATAATAATAATAATAATA 890
 317 GGGAAAGATTAAGACATGATGTTCTTAATAATAATAATAATAATAATAATAATAATA 258
 891 TTATGCGCAGCAAAACCGAATGAGATTCATGTAATAATAATAATAATAATAATAATA 950
 257 TTATGCGCAGCAAAACCGAATGAGATTCATGTAATAATAATAATAATAATAATAATA 198
 951 TGAATTTTATTAAGATGAGGAGCCTTATTTATTTGCTGCTGCTGCTGCTGCTGCTGCT 1010
 197 TGAATTTTATTAAGATGAGGAGCCTTATTTATTTGCTGCTGCTGCTGCTGCTGCTGCT 138
 1011 TACATACATGCTGCTGATGAGGCTA-TTTTTCCTCCATTTAAGAGGATGCAATTA 1069
 137 TACATACATGCTGCTGATGAGGCTA-TTTTTCCTCCATTTAAGAGGATGCAATTA 78
 1070 AATGCTGCTTTTGA 1085
 77 AATGCTGCTTTTGA 62

RESULT 6	530 bp	mrna	linear	EST_06-MAR-2002
LOCUS	BM826722			
DEFINITION	K-EST0098878 S14K402s1 Homo sapiens cDNA clone S14K402s1-16-B01 5',			
ACCESSION	BM826722			
VERSION	BM826722.1			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 530)			
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,T.M., Park,H.S., Kim,S. and Kim,Y.S.			
TITLE	21C Frontier Korean EST Project 2001			
JOURNAL	Unpublished			
COMMENT	Contact: Kim YS			
	Genome Research Center			
	52 Eeann-dong Yuseong-gu, Daejeon 305-333, South Korea			
	Tel: +82-42-860-4470			
	Fax: +82-42-860-4409			
	Email: yongsung@email.krribb.re.kr			
	Plate: 16 row: B column: 01			
	High quality sequence stop: 530.			
FEATURES	Location/Qualifiers			
source	1..530			

BASE COUNT	131 a	88 c	130 g	181 t
ORIGIN				

Query Match	46.38;	Score 507;	DB 12;	Length 530;
Best local Similarity	-99.88;	Pred. No. 3.1e-98;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;

292 TCAGATTTATTTCCCAAGTGATGAAGAAATGAGGGTTCCCATACTTGGGCTGCTG 351

Db 12 TCAGATTATTTCCCAAGTCAATGAAAGAAATAGAGGGTTCCATTAAGCTTTGGGCTGCTG 71

QY 352 CTTGGTTTTTCTCCCTTGCTGCTCGCTCTGTGTGACTCAGGCTGGAAATTTACTGGGGTCA 411

Db 72 CTTGGTTTTTCTCCCTTGCTGCTCGCTGTGTGACTCAGGCTGGAAATTTACTGGGGTCA 131

QY 412 TCGATTGACACCTCTGTGCTGGATGGGGCATTTTAAATTCAGCTATACAGGCTACT 471

Db 132 TCTGATTGACACCTCTGTGCTGGATGGGGCATTTTAAATTCAGCTATACAGGCTACT 191

QY 472 TGGATATATCTGCATTTATGAGAGGAACACATTCATTGAGAGATACAGAAATGATGATGG 531

Db 192 TGGATATATCTGCATTTATGAGAGGAACACATTCATTGAGAGATACAGAAATGATGATGG 251

QY 532 AGCAAGAGAGTGGATATTTCTGGCTATAGTGGAGAGCTTGGCTGTTTGTATTT-ACGGCTA 590

Db 252 AGCAAGAGAGTGGATATTTCTGGCTATAGTGGAGAGCTTGGCTGTTTGTATTTAATTAAGGCTA 311

QY 591 TCCCTTTGATTGCATATTTATTTCTGGTCATTTGGTCATTTGCACATTTCCATAGACCTAATTTATGCG 650

Db 312 TCCCTTTGATTGCATATTTATTTCTGGTCATTTGGTCATTTGCACATTTCCATAGACCTAATTTATGCG 371

QY 651 CAATTTCATACCCCTGACCTGGGGAGTTGCTTTAGGCTGGTGTATGATGATTTGTTTTCGCATTA 710

Db 372 CAATTTCATACCCCTGACCTGGGGAGTTGCTTTAGGCTGGTGTATGATGATTTGTTTTCGCATTA 431

QY 711 TTTGGATTCCAAATTTATGGCTATTCATAAAAAATTAATTCAGGCTAAGGAAACATCTTTCAAC 770

Db 432 TTTGGATTCCAAATTTATGGCTATTCATAAAAAATTAATTCAGGCTAAGGAAACATCTTTCAAC 491

QY 771 GCGTTATAGTGTGCGACAGACAGTTTAACTGAGGCTC 809

Db 492 GCGTTATAGTGTGCGACAGACAGTTTAACTGAGGCTC 530

RESULT 7	BG964290	691 bp	mus musculus	cdna	clone	IMAGE:4983526 5'
LOCUS	BG964290					
DEFINITION	BG964290	NCI-CGAP-Co24	Mus musculus	cdna	clone	IMAGE:4983526 5'
ACCESSION	BG964290					
VERSION	BG964290.1	GI:14351927				
KEYWORDS	EST.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 691)					
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .					
JOURNAL	National Institutes of Health. Mammalian Gene Collection (MGC)					
COMMENT	Unpublished					
	Contact: Robert Strausberg, Ph.D.					
	Email: cgapbs-remail.nih.gov					
	Tissue Procurement: Jeffrey E. Green, M.D.					
	CDNA Library Preparation: Life Technologies, Inc.					
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)					
	DNA Sequencing by: Incyte Genomics, Inc.					
	DNA distribution: MGC clone distribution information can be					
	found through the I.M.A.G.E. Consortium/LNL at:					
	http://image.lnl.gov					
	plate: LLAM10988, row: d column: 23					
	High quality sequence stop: 689.					
	location/Qualifiers					
	1..691					


```

OY 521 ATGATGATTGAGCAAGAGTATTTCTGGCTATGATGATGAGACCTTCTGCTGTTGTA 580
DB 494 ATGATGATTGAGCAAGAGTATTTCTGGCTATGATGATGAGACCTTCTGCTGTTGTA 435
OY 581 ATTAGCCCTATTCCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
DB 434 ATCACTCTATACCTGTTCTGCAATTCCTGATGATGATGATGATGATGATGATGAT 375
OY 641 AATTATGAGCAATTCCTATACCTGATGATGATGATGATGATGATGATGATGATGAT 700
DB 374 AATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 315
OY 701 TTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 760
DB 314 TTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
OY 761 ATCTTTCAGCCCTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 820
DB 254 ATCTTTCAGCCCTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 195
OY 821 CACATGCTGGGGGAAAGATATTAAGACATGCTGATGATGATGATGATGATGATGAT 880
DB 194 AACATGCTGGAGAGAGATATTAAGACATGCTGATGATGATGATGATGATGATGAT 135
OY 881 ATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 937
DB 134 ATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75
OY 938 TGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 997
DB 74 TTCTCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15
OY 998 GAATAGGAAATGT 1011
DB 14 AACTAGGGGATGT 1

RESULT 11
BX523178 standard; RNA; EST; 567 BP.
AC BX523178;
SV BX523178.1
XX
DT 27-MAY-2003 (Rel. 75, Created)
DT 27-MAY-2003 (Rel. 75, Last updated, Version 1)
DE R2PD Mus musculus cDNA clone IMAGP998B202628 - IMAGE:1051315 5' EST.
XX EST: expressed sequence tag.
XX Mus musculus (house mouse)
XX Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
XX Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XX [1]
XX 1-567
XX Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,
XX Korn B.;
XX Submitted (28-MAY-2003) to the EMBL/GenBank/DBJ databases.
XX R2PD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer
XX Feld 580, D-69120 Heidelberg, Germany
XX
XX RZPD: IMAGP998B202628.
XX RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
XX Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
XX http://www.rzpd.de/ClonedCards/cgi-bin/showLib.pl.cgi?response71libNo=981
XX Contact: Ina Rolfs
XX RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
XX Heubnerweg 6, D-14059 Berlin, Germany
XX Tel: +49 30 32639 101

```

```

CC Fax: +49 30 32639 111
CC www.rzpd.de
CC This clone is available royalty-free from RZPD;
CC contact RZPD (clone@rzpd.de) for further information.
CC Seq primer: T7, Primer sequence: TAATACGACTCACTATAGG
XX Key
XX Location/Qualifiers
FH 1..567
FT /db_xref="taxon:10090"
FT /note="This issue obtained from 8 week old mouse. Colon was
FT harvested 72 hours after irradiation with 1400 Gys. 1st
FT strand cDNA was primed with a Not I - oligo (dT) primer
FT (5'-GTGTAAGCAATCGAAGTGGAGGCGGCCGCCCTTTT-3')
FT TTTT-3') double-stranded cDNA was ligated to
FT Eco RI adaptors (AATTCGTGCACTG), digested with NotI and
FT vector. Library constructed by R. Barstead (Oklahoma
FT Medical Research Foundation).
FT /organism="Mus musculus"
FT /clone="IMAGP998B202628"
FT /clone_lib="Barstead irradiated colon MPL-RB7"
FT /dev_stage="8 weeks"
FT /lab_host="DH10B"
SQ
Sequence 567 BP; 158 A; 100 C; 140 G; 169 T; 0 other;
Query Match 37.9%; Score 415; DB 4; Length 567;
Best Local Similarity 86.0%; Pred. No. 1.7e-78;
Matches 460; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
OY 389 CAGCGGGAATTTACTGGGCTCATGATGACCACTTGTGCTGATGGGCAATTTTA 448
DB 26 CAGCGGGAATTTACTGGGCTCATGATGATGATGATGATGATGATGATGATGATGAT 85
OY 449 ATTGCAGCTATGAGCTAGTATGATGATGATGATGATGATGATGATGATGATGAT 508
DB 86 ATTGCAGCTATGAGCTAGTATGATGATGATGATGATGATGATGATGATGATGAT 145
OY 509 GAGGATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
DB 146 GAGGATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 205
OY 569 TGCTGGTTGTATTAACGCTATGCTTTGATGATGATGATGATGATGATGATGATGAT 628
DB 206 TGCTGGTTGTATTAACGCTATGCTTTGATGATGATGATGATGATGATGATGATGAT 265
OY 629 TTTCATAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 688
DB 266 TTTCATAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325
OY 689 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
DB 326 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 385
OY 749 GCTAAAGGAAACATCTTCAACGCTTATTAAGTATGATGATGATGATGATGATGAT 808
DB 386 GCTAAAGGAAACATCTTCAACGCTTATTAAGTATGATGATGATGATGATGATGAT 445
OY 809 CCATACCTGGAACAACTCGTGGGCAAGATATTAAGACATGATGATGATGATGATGAT 868
DB 446 CCATACCTGGAACAACTCGTGGGCAAGATATTAAGACATGATGATGATGATGATGAT 505
OY 869 GCTGCGCATGAATATACCTAGTATGATGATGATGATGATGATGATGATGATGATGAT 923
DB 506 ACTGACCAAGAAATACCTAGTATGATGATGATGATGATGATGATGATGATGATGAT 560

```

```

RESULT 12
CB586665/c 582 bp mRNA linear EST 03-APR-2003
LOCUS AMGNMUC:SRC51-00008-G9-A src51 (10883) Rattus norvegicus cDNA clone
DEFINITION src51-00008-g9 5', mRNA sequence.

```


ACCESSION CB586665.1 GI:29530706
 VERSION
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 582)
 AUTHORS Amgen EST Program.
 TITLE Amgen Rat EST Program
 JOURNAL Unpublished
 COMMENT Contact: Dan Fitzpatrick
 Amgen, Inc
 One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00008 row: 9 column: 9.
 Location/Qualifiers
 1..582
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="src1-00008-99"
 /tissue_type="Cornea"
 /clone_lib="src1 (10883)"
 /note="Vector: PCR4-TOPO; Site_1: ECORI; Site_2: ECORI;
 rat corneas"
 rat corneas"

BASE COUNT 182 a 139 c 92 g 169 t
 ORIGIN
 Query Match 37.6%; Score 411.6; DB 14; Length 582;
 Best Local Similarity 82.8%; Pred. No. 8.9e-78;
 Matches 482; Conservative 0; Mismatches 99; Indels 1; Gaps 1;

398 ATTACTGGGTCATCGATTGACCACTCTGTCGTCGATGGGCAATTTAATTCGAGCT 457
 |||||
 582 ATTACTGGGTCATCGATTGACCACTCTGTCGTCGATGGGCAATTTAATTCGAGCT 523
 |||||
 458 ATTACTGGGTCATCGATTGACCACTCTGTCGTCGATGGGCAATTTAATTCGAGCT 517
 |||||
 522 ATTCTGGAATTCGAGCAATCATCTGATTTATGAGAAATAGTTCATGAGATATA 463
 |||||
 518 GAAATGATTTGGAGCAAGAGATGATTCGCTATGCTGATGGAGAGCTTGGCTTT 577
 |||||
 462 GAAATGATTTGGAGCAAGAGATGATTCGCTATGCTGATGGAGAGCTTGGCTTT 403
 |||||
 578 GTAATTAAGCCATCTTTGATTCGCAATTTATTCGTCATTCGTCGCAATTCATGAGA 637
 |||||
 402 GTCATCAGCTCTATCTGTTGTCGCAATTCGTCGTCGTCGTCGTCGTCGTCGTCG 343
 |||||
 638 CCTAATTAAGGCGCAATTCATCCCTGACTGCGGAGTTCGCTTAGCTGCTGATGATT 697
 |||||
 342 CCTAATTAAGGCGCAATTCATCCCTGACTGCGGAGTTCGCTTAGCTGCTGATGATT 283
 |||||
 698 GTTTCGCAATTTGATTCATCAATTTAGCTATGCTATCAATAAATATTCAGGCTAAAGA 757
 |||||
 282 ATTTCGCAATTTGATTCATCAATTTAGCTATGCTATGCTATCAATAAATATTCAGGCT 223
 |||||
 758 AACATCTTCAAGCCTTTAAGTTCGTCGACAGCAGCTTCACTGAGGCTCCATACCTG 817
 |||||
 222 AACATCTTCAAGCCTTTAAGTTCGTCGACAGCAGCTTCACTGAGGCTCCATACCTG 163
 |||||
 818 GAAACACATGTCGAGGAGATATAAAGACATGTAAGTTCCTAATAAAGAGGCTGGCAT 877
 |||||
 162 GAAACACATGTCGAGGAGATATAAAGACATGTAAGTTCCTAATAAAGAGGCTGGCAT 103
 |||||
 878 GAAATACCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 937
 |||||
 102 GAAATACCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 43
 |||||
 938 TGATTG-TATTAATGTCATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 978
 |||||
 42 GAATGTCCTAATATGTCATTTTGTAGAGTAAAGGAGGGAATTTT 1

RESULT 13
 BE774564
 LOCUS 484 bp mRNA linear EST 25-APR-2001
 DEFINITION 284274 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 BE774564
 BE774564.1 GI:12122464
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bos taurus (cow)
 Bos taurus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 484)
 AUTHORS Smith, P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
 Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett,
 G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-Mckown, C.G.,
 Pette, G., Holt, I., Karaycheva, S., Liang, F., Quackenbush, J. and
 Keefe, J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 11282978
 PUBMED
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -mismatches 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCAGCTGACGAGC
 Plate: 84 row: 9 column: 9
 Seq primer: ATTAGTGACATCATG.
 Location/Qualifiers
 1..484
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 3BOV"
 /note="Vector: pCMV SPOR6; Site_1: NotI; Site_2: SalI;
 Library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."

BASE COUNT 128 a 81 c 116 g 159 t
 ORIGIN
 Query Match 34.8%; Score 381; DB 10; Length 484;
 Best Local Similarity 89.1%; Pred. No. 3.2e-71;
 Matches 434; Conservative 0; Mismatches 50; Indels 3; Gaps 2;

273 AAAGCATCACACAAATTCAGATTTTCCCAAGTATGAGAAATAGAGGCTC 332
 |||||
 1 AAAGCATCACACAGCAATTCAGATTTTCCCAAGTATGAGAAATAGAGGCTC 57
 |||||
 333 CCATTAATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 392
 |||||
 58 CCATTAATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 117
 |||||
 393 CTGGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 452
 |||||
 118 CTGGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 177
 |||||
 453 CAGCTAATCTGAGCTGTTGGAATCATCTGCAATTTATGAGGAGACAGATTCATTGAGG 512
 |||||
 178 CAGCTAATCTGAGCTGTTGGAATCATCTGCAATTTATGAGGAGACAGATTCATTGAGG 237
 |||||

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=4t2-Qv4-BN0090-040
500-199-a06&ts=2000-05-04&ts=1)

Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 360.
Location/Qualifiers

FEATURES

source

1..360

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BN0090"

/note="Organ: breast, normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT

119 a 83 c 61 g 97 t

ORIGIN

Query Match

29.8%; Score 325.8; DB 10; Length 360;

Best Local Similarity 96.4%; Pred. No. 2,1e-59;
Matches 344; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

```
QY 475 AATCATCTGATTTATGAGGAGGAGACATTCATTGAGATACAGAAATGATGGAGC 534
    |||||||
DB 360 AATCATCTGATTTATGAGGAGGAGACATTCATTGAGATACAGAAATGATGGAGC 301
    |||||||
QY 535 AAAGAGTGATTTCTGCTATGCTGAGAGAGCTTCTGTTGTAATTAAGCCTATCT 594
    |||||||
DB 300 AAAGAGTGATTTCTGCTATGCTGAGAGAGCTTCTGTTGTAATTAAGCCTATCT 241
    |||||||
QY 595 TTGATTCGAATTTATCTGCTATGCTGATGCAATTCATTAATGAGCGCAAT 654
    |||||||
DB 240 TTGATTCGAATTTATCTGCTATGCTGATGCAATTCATTAATGAGCGCAAT 181
    |||||||
QY 655 TCCATACCCCTGAGCTGGAGAGTTGCTTTAGGCTGGTGTATGATGTTTCTGCAATATTG 714
    |||||||
DB 180 TCCATACCCCTGAGCTGGAGAGTTGCTTTAGGCTGGTGTATGATGTTTCTGCAATATTG 121
    |||||||
QY 715 GATTCGAATTTATGCTATTCATTAATAATTAATCAGGCTAAAGAAACATCTTCAAGCGCT 774
    |||||||
DB 120 GATTCGAATTTATGCTATTCATTAATAATTAATCAGGCTAAAGAAACATCTTCAAGCGCT 61
    |||||||
QY 775 TATAGTTCGTCAGACGAGTTCTTAACCTGGGCTCATATACCTGGAACAACATCGTGG 831
    |||||||
DB 60 TATAGTTCGTCAGACGAGTTCTTAACCTGGGCTCATATACCTGGAACAACATCGTGG 5
```

Search completed: August 27, 2003, 12:07:17
Job time : 2784 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 17:32:44 ; Search time 29 Seconds

(without alignments)
315.143 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198
Sequence: 1 SEETITTTIQLDFPKVKKMVPRKEAGHEIPTVSGSRKPE 216

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	504.5	42.1	799 2	US-08-700-013B-27 Sequence 27, Appl
2	503.5	42.0	797 3	US-09-182-728A-2 Sequence 2, Appl
3	503.5	42.0	797 4	US-09-191-468-120 Sequence 120, App
4	503.5	42.0	797 4	US-09-191-468-122 Sequence 122, App
5	503.5	42.0	797 4	US-09-795-232-2 Sequence 2, Appl
6	502.5	41.9	280 2	US-08-700-013B-17 Sequence 17, Appl
7	502.5	41.9	797 2	US-08-700-013B-19 Sequence 19, Appl
8	502.5	41.9	797 2	US-08-700-013B-15 Sequence 15, Appl
9	501.5	41.9	280 2	US-08-700-013B-21 Sequence 21, Appl
10	501.5	41.9	797 2	US-08-700-013B-25 Sequence 25, Appl
11	462.5	37.4	194 2	US-08-700-013B-23 Sequence 23, Appl
12	447.5	37.4	194 2	US-08-700-013B-25 Sequence 25, Appl
13	359	30.0	667 1	US-07-879-617A-8 Sequence 8, Appl
14	359	30.0	667 1	US-08-753-985-8 Sequence 8, Appl
15	350.5	29.3	150 4	US-09-191-468-88 Sequence 92, Appl
16	350.5	29.3	150 4	US-09-191-468-92 Sequence 92, Appl
17	350.5	29.3	150 4	US-09-191-468-94 Sequence 94, Appl
18	350.5	29.3	150 4	US-09-191-468-96 Sequence 96, Appl
19	350.5	29.3	150 4	US-09-191-468-98 Sequence 98, Appl
20	350.5	29.3	150 4	US-09-191-468-100 Sequence 100, App
21	350.5	29.3	150 4	US-09-191-468-102 Sequence 102, App
22	345.5	28.8	687 3	US-08-834-467-2 Sequence 2, Appl
23	345.5	28.8	687 3	US-09-396-177-2 Sequence 2, Appl
24	344.5	28.8	150 4	US-09-191-468-90 Sequence 90, Appl
25	344.5	28.8	638 1	US-08-295-814E-13 Sequence 13, Appl
26	344.5	28.8	638 1	US-08-240-783B-2 Sequence 2, Appl
27	344.5	28.8	638 3	US-09-084-813-2 Sequence 2, Appl

28	344.5	28.8	638 3	US-09-343-361-13 Sequence 13, Appl
29	344.5	28.8	638 5	PCT-US92-09662-2 Sequence 2, Appl
30	341	28.5	635 1	US-07-879-617A-9 Sequence 9, Appl
31	341	28.5	635 1	US-08-753-985-9 Sequence 9, Appl
32	325	27.1	616 1	US-07-879-617A-13 Sequence 13, Appl
33	325	27.1	616 1	US-08-753-985-13 Sequence 13, Appl
34	319	26.6	620 1	US-08-301-722A-2 Sequence 2, Appl
35	316	26.4	619 1	US-07-762-132A-2 Sequence 2, Appl
36	316	26.4	619 1	US-08-301-722A-2 Sequence 2, Appl
37	308.5	25.8	599 1	US-07-879-617A-10 Sequence 10, Appl
38	308.5	25.8	599 1	US-08-295-814E-11 Sequence 11, Appl
39	308.5	25.8	599 1	US-08-240-783B-4 Sequence 4, Appl
40	308.5	25.8	599 3	US-08-753-985-10 Sequence 10, Appl
41	308.5	25.8	599 3	US-09-084-813-4 Sequence 4, Appl
42	308.5	25.8	599 3	US-09-343-361-11 Sequence 11, Appl
43	308.5	25.8	599 5	PCT-US92-09662-4 Sequence 4, Appl
44	299	25.0	617 1	US-07-879-617A-11 Sequence 11, Appl
45	299	25.0	617 1	US-08-301-722A-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-700-013B-27
Sequence 27, Application US/08700013B
Patent No. 5919653
GENERAL INFORMATION:
APPLICANT: Albert, Vivian R.
APPLICANT: Kowalski, Leslie R.
APPLICANT: Borden, Laurence A.
APPLICANT: McKelvey, Jeffrey F.
TITLE OF INVENTION: Human Glycine Transporter
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,013B
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER: 317743-108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 799 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-700-013B-27
Query Match
Best Local Similarity 42.1%; Score 504.5; DB 2; Length 799;
Matches 91; Conservative 33; Mismatches 65; Indels 3; Gaps 1;


```

? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Bloom, Allen
? REGISTRATION NUMBER: 29,135
? REFERENCE/DOCKET NUMBER: 317743-108
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 609-520-3214
? TELEFAX: 609-520-3259
? TELEX:
? INFORMATION FOR SEQ ID NO: 17:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 280 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? US-08-700-013B-17

```

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

? REGISTRATION NUMBER: 29,135
? REFERENCE/DOCKET NUMBER: 317743-108
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 609-520-3214
? TELEFAX: 609-520-3259
? TELEX:
? INFORMATION FOR SEQ ID NO: 19:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 797 amino acids
? type: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
US-08-700-013b-19

Query Match	41.98;	Score	503.5;	DB	2;	Length	797;
Best Local Similarity	46.48;	Pred.	No. 1.8e-46;				
Matches	91;	Conservative	33;	Mismatches	69;	Indels	3;
						Gaps	1.

[illegible]

```

RESULT 8
US-09-191-468-124
? Sequence 124, Application US/09191468A
? Patent No. 6416975
? GENERAL INFORMATION:
? APPLICANT: Gallagher, Michael J.
? APPLICANT: Burgess, Loyd R.
? APPLICANT: Brunden, Kurt R.
? TITLE OF INVENTION: Human Glycine Transporter Type 2
? FILE REFERENCE: 12311US01
? CURRENT APPLICATION NUMBER: US/09/191,468A
? CURRENT FILING DATE: 1998-11-12
? NUMBER OF SEQ ID NOS: 124
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 124
? LENGTH: 797
? TYPE: PRT
? ORGANISM: Human
US-09-191-468-124

```

Query Match	41.9%	Score 502.5;	DB 4;	Length 797;
Best Local Similarity	46.4%	Pred. No. 1.8e-46;		
Matches 91; Conservative	33;	Mismatches 69;	Indels 3;	Gaps 1

```

QY 3 EITTTIODLFPKXKKKRVPTLSCCLVFLSLGVCVQASAIYWHLLDHCACGAGSILI 62
Dy 584 ETIVISIDEPFKYLRTHKPVETLSCCLCEFTMGPPMIIQGGIYMFQLDVTYAAASALVI 643
QY 63 AALLETVGLIYGGNRFIEDREMMIGAKRWLEWIMWRACQWVPIPLIAIFMSLYOF 122
Dy 644 IAIPELVGISYVYQGRFCEDIEMMIGQPNIF--WKVCMAFVPTILTFLTCSEFYGM 700
QY 123 HRPNNGAIPYDPDMGVALGCMIVFCIIWIPIWAATIKIQAKNINIPORLLSCCRPASNMGP 182
Dy 701 EPMYTGSRYPNMSMWLGLMLACSVIWIPIMEVIMKLHAPGRIFERLEKLVCSPPDPMGP 760
QY 183 YLEOHRGERYKDMVVP 198

```

Db 761 FLAQHGRERYKNMIDP 776

RESULT 9
 US-08-700-013B-15
 Sequence 15, Application US/08700013B
 Patent No. 5919653
 GENERAL INFORMATION:
 APPLICANT: Albert, Vivian R.
 APPLICANT: Kowalski, Leslie R.Z.
 APPLICANT: Borden, Laurence A.
 APPLICANT: McKeivly, Jeffrey F.
 TITLE OF INVENTION: Human Glycine Transporter
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dechart Price & Rhoads
 STREET: 997 lenox Drive, Building 3, Suite 2100
 CITY: Lawrenceville
 STATE: NJ
 COUNTRY: USA
 ZIP: 08543
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: Fastseq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/700, 013B
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Bloom, Allen
 REGISTRATION NUMBER: 29,135
 REFERENCE/DOCKET NUMBER: 317743-108
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 609-520-3214
 TELEFAX: 609-520-3259
 TELEX:
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 280 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-700-013B-15

Query Match	41.9%	Score	501.5;	DB 2;	Length	280;			
Best Local Similarity	46.4%	Pred. No.	6.4e-47;						
Matches	91;	Conservative	33;	Mismatches	69;	Indels	3;	Gaps	1;

[illegible]

```

? TITLE OF INVENTION: Human Glycine Transporter
? NUMBER OF SEQUENCES: 41
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Dechert Price & Rhoads
? STREET: 997 Lenox Drive, Building 3, Suite 210
? CITY: Lawrenceville
? STATE: NJ
? COUNTRY: USA
? ZIP: 08543
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/700,013B
? FILING DATE:
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Bloom, Allen
? REGISTRATION NUMBER: 29,135
? REFERENCE/DOCKET NUMBER: 317743-108
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 609-520-3214
? TELEFAX: 609-520-3259
? TELEX:
? INFORMATION FOR SEQ ID NO: 25:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 194 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? US-08-700-013B-25

Query Match      38.6%; Score 462.5; DB 2; Length 194;
Best Local Similarity 47.1%; Pred. No. 7.5e-43;
Matches 82; Conservative 30; Mismatches 59; Indels 3; Gaps 1

QY 25 TLGGCLVFLGLGVCTVAGIYWHNLIDHFCAGMCIILAIILELVIIGTIWVGNNFIEDT 84
||||| :||| :||| ||| :| :| :| | ||| :||| |||
Db 3 TLGGCICFFIMGFPMITGGIYMQLVPTVAASVALVIAIFELGISYVVGLGRFCBDI 62
||||| :||| :||| ||| :| :| :| | ||| :||| |||
QY 85 EMMIGAKRMIFMLWMRACMFVTIPTLLAIFATWSLVOPHRDNYGALPRPDGMALMGOMI 144
||||| :||| :||| ||| :| :| :| | ||| :||| |||
Db 63 EMMIGFOGNIF--WKVCMAFVTPILLFLIFLCFSFYQNEPMTYSRYKRPNSMWLGWLML 119
||||| :||| :||| ||| :| :| :| | ||| :||| |||
QY 145 VFCIIIPDIALIITIIIOAKNIFIORLLISCCRPASNMGPELEOHNGERYKDMVP 198
||||| :||| :||| ||| :| :| :| | ||| :||| |||
Db 120 ACSVIMPIKMSVKIRMHLPGRFIERLKLYVCSPPDPDWGPFQAQHGERYKNMDP 173
||||| :||| :||| ||| :| :| :| | ||| :||| |||

RESULT 12
US-08-700-013B-23
Sequence 23, Application US/08700013B
Patent No. 5919653
GENERAL INFORMATION:
APPLICANT: Albert, Vivian R.
APPLICANT: Kowalski, Leslie R.Z.
APPLICANT: Borden, Laurence A.
APPLICANT: McKelvy, Jeffrey F.
TITLE OF INVENTION: Human Glycine Transporter
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,013B
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER: 317743-108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-700-013B-23

Query Match 37.48; Score 447.5; DB 2; Length 194;
Best Local Similarity 46.08; Pred. No. 3.3e-41;
Matches 80; Conservative 30; Mismatches 61; Indels 3; Gaps 1;
QY 25 TLGCLVLFLLGLVCTOAGIYVHLIDHCAGKGLIIAILELVGIMVIGGNRIEDT 84
DB 3 TLGCIYCFIFGFMPIYGGIYMFQIDYTAASAYALVILAFELVGIISYVGLQRCEDI 62
QY 85 EMMTGARWTFWLMWRACWPIYPIILIAIFIMSLVOFHPRNGAIPYPMGVALGCM 144
DB 63 EMMIGFQNI--WKQCAVPTITITFLICFSFYQWEMXTYGSYPRMMSVGLGLML 119
QY 145 VECIWIPIMAIKIITQAKNIFORLISCCRPASNMGPIYEQHNGEKIKMVP 198
DB 120 ACSYIPIPIVVMKHLAPGRIFIRLKLVCSPQPMGPFLAORHGERYKNWIDP 173

RESULT 13
US-07-879-617A-8
Sequence 8, Application US/07879617A
Patent No. 5580775
GENERAL INFORMATION:
APPLICANT: Fremeanu Jr., Robert T.
APPLICANT: Caron, Marc G.
APPLICANT: Blakey, Randy D.
TITLE OF INVENTION: A High Affinity L-Proline Transporter
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
SERRET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/879,617A
FILING DATE: 19920501
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Padst, Patricia L.
REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: EMU109
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 667 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Rattus
TISSUE TYPE: Brain
IMMEDIATE SOURCE:
LIBRARY: rat forebrain cDNA library
CLONE: RTB2-2-20
FEATURE:
NAME/KEY: Domain
LOCATION: 76..95
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 103..127
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 147..167
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 247..266
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 272..294
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 321..341
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 353..375
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 404..427
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 454..473
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 487..509
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 530..549
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 567..589
OTHER INFORMATION: /note= "Membrane-spanning domain"
FEATURE:
NAME/KEY: Region
LOCATION: 44..45
OTHER INFORMATION: /note= "protein kinase C phosphorylation site"
OTHER INFORMATION:

Query Match	30.0%;	Score 359;	DB 1;	Length 667;
Best Local Similarity	30.4%;	Pred. No. 6.96-31;		
Matches	65;	Conservative 55;	Mismatches 78;	Indels 16; Gaps 2
OY	3	ETITTTTIODLFFPKVKKMKRVPTLGCCLVLELGVVQAGIYVWHLIDHRCAGGILTI	62	
Db	435	ETITVTAVTDEFPYLRPKKAVFSGILCAVMTLMLITTTGGMTVLVLLDDTSASFGLMV	494	
OY	63	AATLELVGIMLYGSGNRFIETEMMIGAKRRIEMLRACVFETPILITIAIFISLVOF	122	
Db	495	VYITTCVLAVTRYVGIORFCRDIHMMLEGPK---GLYFRACWELTSPATLLALLVSIKY	551	
OY	123	HRPNTGALPRPDGVALGCMIVFCIMIPYIMAKITIOAKNIPRILSCCRPASMNGP	182	
Db	552	QPSSEGVSYFFPAMLELLGILMGILSCIMIPAGMLVAVLEBGSLSMERLQASRPADWGP	611	
OY	183	YLEHGRGERHYDMVYPRKKEGHELPYVSGSKAPE	216	
Db	612	SLEENR-----TGMTVAITLQASQSPK	632	

RESULT 14
US-08-753-985-8
Sequence 8, Application US/08753985
Patent No. 5758788
GENERAL INFORMATION:
APPLICANT: Fremeau Jr., Robert T.
APPLICANT: Caron, Marc G.
APPLICANT: Blakely, Randy D.
TITLE OF INVENTION: A High Affinity L-Proline Transporter
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/753,985
FILING DATE: 03-DEC-1996
CLASSIFICATION: 536

```

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: US 07/879617
3 FILING DATE: 01-MAY-1992
4 ATTORNEY/AGENT INFORMATION:
5 NAME: Pabst, Patricia L.
6 REGISTRATION NUMBER: 31,284
7 REFERENCE/DOCKET NUMBER: EMT09
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: 404-815-6508
10 TELEFAX: 404-815-6555
11 INFORMATION FOR SEQ ID NO: 8:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 667 amino acids
14 TYPE: amino acid
15 STRANDEDNESS: single
16 TOPOLOGY: linear
17 MOLECULE TYPE: protein
18 HYPOTHEetical: NO
19 ANTI-SENSE: NO
20 FRAGMENT TYPE: N-terminal
21 ORIGINAL SOURCE:
22 ORGANISM: Rattus
23 TISSUE TYPE: Brain
24 IMMEDIATE SOURCE:
25 LIBRARY: rat forebrain cDNA library
26 CLONE: ITB2-2-20
27 FEATURE:
28 NAME/KEY: Domain
29 LOCATION: 76..95
30 OTHER INFORMATION: /note= "Membrane-spanning domain"
31 FEATURE:
32 NAME/KEY: Domain
33 LOCATION: 103..127
34 OTHER INFORMATION: /note= "Membrane-spanning domain"
35 FEATURE:
36 NAME/KEY: Domain
37 LOCATION: 147..167
38 OTHER INFORMATION: /note= "Membrane-spanning domain"
39 FEATURE:
40 NAME/KEY: Domain
41 LOCATION: 247..266
42 OTHER INFORMATION: /note= "Membrane-spanning domain"
43 FEATURE:
44 NAME/KEY: Domain
45 LOCATION: 272..294
46 OTHER INFORMATION: /note= "Membrane-spanning domain"
47 FEATURE:
48 NAME/KEY: Domain
49 LOCATION: 321..341
50 OTHER INFORMATION: /note= "Membrane-spanning domain"
51 FEATURE:
52 NAME/KEY: Domain
53 LOCATION: 353..375
54 OTHER INFORMATION: /note= "Membrane-spanning domain"
55 FEATURE:
56 NAME/KEY: Domain
57 LOCATION: 404..427
58 OTHER INFORMATION: /note= "Membrane-spanning domain"
59 FEATURE:
60 NAME/KEY: Domain
61 LOCATION: 454..473
62 OTHER INFORMATION: /note= "Membrane-spanning domain"
63 FEATURE:
64 NAME/KEY: Domain
65 LOCATION: 487..509
66 OTHER INFORMATION: /note= "Membrane-spanning domain"
67 FEATURE:
68 NAME/KEY: Domain
69 LOCATION: 530..549
70 OTHER INFORMATION: /note= "Membrane-spanning domain"
71 FEATURE:
72 NAME/KEY: Domain
73 LOCATION: 567..589

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 17:36:34 ; Search time 52 Seconds

(without alignments)
568.171 Million cell updates/sec

Title: US-09-940-227-83

Perfect score: 1198
Sequence: 1 SEETITTTQDLEPKVKMKM.....VPKKEAGHEIPVSGSRKPE 216

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 510680 seqs, 136781880 residues

Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubppa/PCCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppa/PCRTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	503.5	42.0	797	9	US-09-795-232-2
2	361	30.1	636	12	US-10-289-161A-3
3	344.5	28.8	633	12	US-10-205-219-30
4	344.5	28.8	638	12	US-09-845-908-13
5	319	26.6	620	12	US-10-241-220-109
6	308.5	25.8	599	12	US-09-845-908-11
7	296	24.7	599	12	US-10-345-680-5
8	295.5	24.7	624	9	US-09-795-693-24
9	295.5	24.7	624	15	US-10-156-238-24
10	295.5	24.7	624	15	US-10-199-485-24
11	287.5	24.0	727	10	US-09-815-923-10
12	286.5	23.9	627	12	US-09-845-908-4
13	283.5	23.7	602	12	US-09-845-908-2
14	281.5	23.5	587	10	US-09-815-923-4
15	280	23.4	622	9	US-09-843-598-11

16	279.5	23.3	246	12	US-09-845-908-8	Sequence 8, Appl1
17	277.5	23.2	632	12	US-09-845-908-12	Sequence 10, Appl1
18	273	22.8	614	12	US-09-845-908-10	Sequence 12, Appl1
19	264	22.0	614	11	US-09-919-039-378	Sequence 378, Appl
20	255.5	21.3	621	12	US-09-845-908-6	Sequence 6, Appl1
21	251.5	21.0	635	12	US-10-205-194-33	Sequence 33, Appl1
22	248	20.7	633	12	US-10-298-974-2	Sequence 2, Appl1
23	248	20.7	633	12	US-09-991-658-2	Sequence 2, Appl1
24	246	20.5	597	10	US-09-815-923-14	Sequence 14, Appl1
25	243.5	20.3	630	9	US-09-843-598-10	Sequence 10, Appl1
26	242.5	20.2	730	9	US-09-741-149-2	Sequence 2, Appl1
27	242.5	20.2	730	9	US-09-795-693-5	Sequence 5, Appl1
28	242.5	20.2	730	12	US-10-385-614-2	Sequence 2, Appl1
29	242.5	20.2	730	15	US-10-156-239-5	Sequence 5, Appl1
30	242.5	20.2	730	15	US-10-156-239-5	Sequence 5, Appl1
31	235.5	19.7	729	9	US-09-741-149-4	Sequence 4, Appl1
32	235.5	19.7	729	12	US-10-385-614-4	Sequence 4, Appl1
33	227.5	18.6	610	10	US-09-861-846-2	Sequence 2, Appl1
34	222	18.5	599	10	US-09-815-923-12	Sequence 2, Appl1
35	210	17.5	599	10	US-09-861-846-4	Sequence 4, Appl1
36	209.5	17.5	727	9	US-09-923-444A-2	Sequence 2, Appl1
37	209.5	17.5	556	10	US-09-815-923-6	Sequence 6, Appl1
38	208	17.4	671	9	US-09-843-598-5	Sequence 5, Appl1
39	200	16.7	671	9	US-09-843-598-7	Sequence 7, Appl1
40	131	10.9	196	11	US-09-989-442-87	Sequence 87, Appl1
41	104	8.7	454	15	US-10-156-761-12378	Sequence 12378, A
42	100.5	8.4	579	10	US-09-738-626-4648	Sequence 1066, Ap
43	92	7.7	188	12	US-10-017-161-1046	Sequence 6706, Ap
44	89	7.4	463	10	US-09-738-626-6706	Sequence 56, Appl1
45	88.5	7.4	501	10	US-09-934-866-56	

ALIGNMENTS

RESULT 1
US-09-795-232-2
Sequence 2, Application US/09795232
Patent No. US20010012627A1
GENERAL INFORMATION:
APPLICANT: Anthony M. Brown
APPLICANT: Conrad Gerald Chapman
APPLICANT: Israel Simon Gloger
APPLICANT: Joanne Rachel Evans
APPLICANT: William Cairns
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP-30176-D1
CURRENT APPLICATION NUMBER: US/09/795,232
PRIORITY FILING DATE: 2001-02-28
PRIORITY APPLICATION NUMBER: 09/182,728
PRIORITY FILING DATE: 1998-10-29
PRIORITY APPLICATION NUMBER: 9818890.7
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 797
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-795-232-2
Query Match
Best Local Similarity 42.0%: Score 503.5; DB 9; Length 797;
Matches 91; Conservative 33; Mismatches 69; Indels 3; Gaps 1;
OY ETITTTQDLEPKVKMKRVPITLGGCLVFLGLVCVDTAGIYVWHLIDHFCAGGIIT 62
DB ETITTSIDSEPKLRTKHFVFTLGGCICFFHGPMTGGIYMQVDVTAASALVI 643
OY AAILELGIWIGNGNFIEDTEMIGAKRWIMMMWRACWPIITPILLIAIFMSLVQF 122
DB IAIPELVGISYVVGLOKFCEDIMMIGFQPNIE---WKVCMAEVPITLITLFCFSFYQM 700

Publication No. US200301484
GENERAL INFORMATION:
APPLICANT: Frantz, Gretchen

```

APPLICANT: Hillan, Kenneth J.
APPLICANT: Phillips, Heidi
APPLICANT: Polakis, Paul
APPLICANT: Spencer, Susan
APPLICANT: Williams, P. Mickey
APPLICANT: Wu, Thomas
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
FILE REFERENCE: P5010R1-US
CURRENT APPLICATION NUMBER: US/10/241,220
CURRENT FILING DATE: 2002-12-13
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 109
LENGTH: 620
TYPE: PRT
ORGANISM: Homo Sapien
US-10-241-220-109

```

```

Query Match
Best Local Similarity 26.6%; Score 319; DB 12; Length 620;
Matches 63; Conservative 32; Mismatches 75; Indels 4; Gaps 2;

```

```

QY 3 ETITTTIDLEPKVKMKRVPTTGGCLVFLGLVCTVQAGIYVHLLIDHFCAGWGIL 62
DB 428 EGVITGLDEFQLHRRRELFTEFLVATFLSLFCVTNGCIYVFTLLDHFAGTSLF 486
QY 63 AAILELVGIIYGGNRFIEDTEMIGAKRWIFWLMWRACFEVTPILLIAFIWSLVQF 122
DB 487 GVLIAIGVAMFYGVGFSDDIQMTGGRP---SLVRLCKKLVSDFLEVVVSVITVF 543
QY 123 HRPNYGAIPYDPMGVAGMCMIVFCIIWPIMAIKITIOAKNIFORLISCRP 176
DB 544 RPHHGAIFPDMANALGWSVATISSMAWVPIYAAYKCSLPSGRKRLAYAIAP 597

```

```

RESULT 6
US-09-845-908-11
; Sequence 11, Application US/09845908
; Publication No. US20030143729A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Kelli E. et al
; TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
; FILE REFERENCE: 40558-D
; CURRENT APPLICATION NUMBER: US/09/845,908
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 11
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-845-908-11

```

```

Query Match
Best Local Similarity 25.8%; Score 308.5; DB 12; Length 599;
Matches 67; Conservative 38; Mismatches 89; Indels 21; Gaps 4;

```

```

QY 3 ETITTTIDLEPKVKMKRVPTTGGCLVFLGLVCTVQAGIYVHLLIDHFCAGWGIL 61
DB 402 EGVITGLDEFQLHRRRELFTEFLVATFLSLFCVTNGCIYVFTLLDHFAGTSLF 461
QY 62 AAILELVGIIYGGNRFIEDTEMIGAKRWIFWLMWRACFEVTPILLIAFIWSLVQ 121
DB 462 FLVFECSVSIWYGVRFNDIQEMVGRPC---WMKLCMSFETPIIVAGVLFSAVQ 518
QY 122 FHRPNYGAIPYDPMGVAGMCMIVFCIIWPIMAIKITIOAKNIFORLISCRPASNG 181
DB 519 MTPILMGSVYFPMKGGVGMALSSWVLIPGMATFALKSLKORIQVNVOPSD 574
QY 182 PYLEHREBRYKDMVYPRKKEGHEIPTVSGSRKPE 216

```

```

DB 575 -----EDIVRP--ENGPRQAGSSASKE 596

```

```

RESULT 7
US-10-345-680-5
; Sequence 5, Application US/10345680
; Publication No. US20030148394A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Slios-Santiago, Immaculada
; APPLICANT: Venkateswarlu, Karicheti
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,
FILE REFERENCE: 641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,
CURRENT APPLICATION NUMBER: US/10/345,680
CURRENT FILING DATE: 2003-01-16
PRIOR APPLICATION NUMBER: US 60/349,511
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/360,500
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/365,041
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/374,063
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/403,468
PRIOR FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: US 60/414,262
PRIOR FILING DATE: 2002-09-27
PRIOR APPLICATION NUMBER: US 60/419,986
PRIOR FILING DATE: 2002-10-21
PRIOR APPLICATION NUMBER: US 60/423,809
PRIOR FILING DATE: 2002-11-05
PRIOR APPLICATION NUMBER: US 60/429,797
PRIOR FILING DATE: 2002-11-26
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 599
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-345-680-5

```

```

Query Match
Best Local Similarity 24.7%; Score 296; DB 12; Length 599;
Matches 57; Conservative 37; Mismatches 80; Indels 4; Gaps 2;

```

```

QY 3 ETITTTIDLEPKVKMKRVPTTGGCLVFLGLVCTVQAGIYVHLLIDHFCAGWGIL 61
DB 402 EGVITGLDEFQLHRRRELFTEFLVATFLSLFCVTNGCIYVFTLLDHFAGTSLF 461
QY 62 AAILELVGIIYGGNRFIEDTEMIGAKRWIFWLMWRACFEVTPILLIAFIWSLVQ 121
DB 462 FLVFECSVSIWYGVRFNDIQEMVGRPC---WMKLCMSFETPIIVAGVLFSAVQ 518
QY 122 FHRPNYGAIPYDPMGVAGMCMIVFCIIWPIMAIKITIOAKNIFORLISCRPASN 179
DB 519 MTPILMGVYFPMKGGVGMALSSWVLIPGMATFALKSLKORIQVNVOPSD 576

```

```

RESULT 8
US-09-795-693-24
; Sequence 24, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 32613, No. US20020068710A1 Human Transporters
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906

```

Best Local Similarity	32.35%	FREQ. NO. 1100 227

APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A1
FILE REFERENCE: 023070-093800US
CURRENT APPLICATION NUMBER: US/09/815,923
CURRENT FILING DATE: 2001-03-23
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 727
TYPE: PRT
ORGANISM: Manduca sexta
FEATURE:
OTHER INFORMATION: neurotransmitter transporter encoded by inbred rat
US-09-815-923-10

Query Match 24.0%; Score 287.5; DB 10; Length 727;
Best Local Similarity 35.0%; Pred. No. 1.1e-20;
Matches 63; Conservative 36; Mismatches 74; Indels 7; Gaps 3;

QY 3 ETTITTTQDLFPKVMKRV---PITLGGCLVFLGLGVCVTOAGIYVWHLIDHFCAGWG 59
Db EVVVTSIDGFPDMIRKRLVHLLVLCVAVSLGCPHIIHSGIYFQMDYVAAASLS 584
QY 60 ILIAAILELVGIWIYGNRFIEDTEMIGAKRWIFWLMWRCWVITPILIAFIWLSL 119
Db ITLAFEEVVAIAWFGVGRSLRNITKMTGRQP---SLYFRCCWIASPALLALMWASM 641
QY 120 VOFHRRNYGAIPIYPMGVALGCMIVFCIIPMAIKIIQKANG-NIFORLISCRPAS 178
Db 642 VDTPEPSYRQYQYPAWMAQLGWMASLSLSCIPYAVAVIIRABDSLRKRLRSIOPTS 701

RESULT 12
US-09-845-908-4
Sequence 4, Application US/09845908
Publication No. US20030143729A1
GENERAL INFORMATION:
APPLICANT: Smith, Kelli E. et al
TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
FILE REFERENCE: 40558-D
CURRENT APPLICATION NUMBER: US/09/845,908
CURRENT FILING DATE: 2001-04-30
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.0 - beta
SEQ ID NO 4
LENGTH: 627
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-845-908-4

Query Match 23.9%; Score 286.5; DB 12; Length 627;
Best Local Similarity 34.3%; Pred. No. 1.2e-20;
Matches 62; Conservative 39; Mismatches 73; Indels 7; Gaps 4;

QY 3 ETTITTTQDLFPKVMKRV---PITLGGCLVFLGLGVCVTOAGIYVWHLIDHFCAGWG 59
Db EVVVTSIDGFPDMIRKRLVHLLVLCVAVSLGCPHIIHSGIYFQMDYVAAASLS 470
QY 60 ILIAAILELVGIWIYGNRFIEDTEMIGAKRWIFWLMWRCWVITPILIAFIWLSL 119
Db 471 ILTVAIFECVCIWVGSNRFYDIEDMIGYRFLSLIK---CMKVYFPGICAGIFIFPL 527
QY 120 VOFHRRNYGAIPIYPMGVALGCMIVFCIIPMAIKIIQKANG-NIFORLISCRPAS 178
Db 528 VTKKPKLVNYYVTPAMYGIGIMLALSSMLCIPIMIFIKLWKTBEIPKQLKLTVBSA 587
QY 179 N 179
Db 588 D 588

RESULT 13
US-09-845-908-2
Sequence 2, Application US/09845908
Publication No. US20030143729A1
GENERAL INFORMATION:
APPLICANT: Smith, Kelli E. et al
TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
FILE REFERENCE: 40558-D
CURRENT APPLICATION NUMBER: US/09/845,908
CURRENT FILING DATE: 2001-04-30
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.0 - beta
SEQ ID NO 2
LENGTH: 602
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-845-908-2

Query Match 23.7%; Score 283.5; DB 12; Length 602;
Best Local Similarity 33.7%; Pred. No. 2.3e-20;
Matches 61; Conservative 40; Mismatches 73; Indels 7; Gaps 4;

QY 3 ETTITTTQDLFPKVMKRV---PITLGGCLVFLGLGVCVTOAGIYVWHLIDHFCAGWG 59
Db EVVVTALVDMPVPRFKRNRREILIVSVSFFGLIMLTGGMVFFQDFYVAAASGC 455
QY 60 ILIAAILELVGIWIYGNRFIEDTEMIGAKRWIFWLMWRCWVITPILIAFIWLSL 119
Db 456 ILTVAIFESLCVAVWYGARFNDIEDMIGYRFLP---IKTCWLFPPACVLAFFLSL 512
QY 120 VOFHRRNYGAIPIYPMGVALGCMIVFCIIPMAIKIIQKANG-NIFORLISCRPAS 178
Db 513 IKTPPLTKKYYTPWVGALGWLALSSMVICIPAMSIYKLTIKGPRERLRQLVCRAE 572
QY 179 N 179
Db 573 D 573

RESULT 14
US-09-815-923-4
Sequence 4, Application US/09815923
Publication No. US20020197644A1
GENERAL INFORMATION:
APPLICANT: Gill, Sarjeet S.
APPLICANT: Ross, Linda S.
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A
FILE REFERENCE: 023070-093800US
CURRENT APPLICATION NUMBER: US/09/815,923
CURRENT FILING DATE: 2001-03-23
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 587
TYPE: PRT
ORGANISM: Manduca sexta
FEATURE:
OTHER INFORMATION: serotonin transporter
US-09-815-923-4

Query Match 23.5%; Score 281.5; DB 10; Length 587;
Best Local Similarity 30.9%; Pred. No. 3.6e-20;
Matches 54; Conservative 34; Mismatches 82; Indels 5; Gaps 2;

QY 3 ETTITTTQDLFPKVMKRV---PITLGGCLVFLGLGVCVTOAGIYVWHLIDHFCAGWG 62
Db 402 EAVTTALCDEYPRVLGRHREVEFAVALLFLFYICALPTTYGGVIVLVDLNVGRLAIFL 461
QY 63 AAILELVGIWIYGNRFIEDTEMIG-AKRWIFWLMWRCWVITPILIAFIWLSLQV 121

```

Db      462 VVFAEAGVCWYGVDRSEEDVRLMIGHPGW----FWRTCKSYISPVFELLVLFVFSVLA 517
OY      122 FHRPNYGAIPYDPWGVALGCMIVFCIIMIPMAIKIIOAKGNIFORLISCCRP 176
        518 HEMELGCEYTPSWSTITVGVMTGTIVSCIPLYIYIKLITPGNCINRIKTIORP 572

```

RESULT 15

```

US-09-843-598-11
; Sequence 11, Application US/09843598
; Patent No. US2002001094A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Ranganathan, Rajesh
; TITLE OF INVENTION: CDSERT GENES, PROTEINS, AND MODULATORY
; FILE REFERENCE: 01997/525002
; CURRENT APPLICATION NUMBER: US/09/843,598
; PRIOR FILING DATE: 2001-04-26
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-843-598-11

```

Query Match 23.4%; Score 280; DB 9; Length 622;

Best Local Similarity 31.9%; Pred. No. 5.5e-20; Mismatches 90; Indels 4; Gaps 2;

```

OY      3 ETITTTIDLPKXMKKRVPTITGCCIVFLGIVCVTOAGIYVHLIDHFCAGWGILI 62
        435 EMATIALCDEPRVYGRRELFLLLAFILCALPTMTYGVVLFVNLVYGPGLALIF 494
OY      63 AAILLVGIIWYGNRPEDTEMMIGAKRMIFWLMWRACWFTVTPILLAIPTMSLVQF 122
        495 VVFEAGVFWYGVDRFSSDVEQMLGSKPGLF--WRICWTYISPVFLITFLFSTMGY 551
OY      123 HRPNTGAIPYDPWGVALGCMIVFCIIMIPMAIKIIOA-KGNIFORLISCCRPASNMG 181
        552 KEMLCERYYPDMYSTQGVANVICSVCIPMTIITKFFFAKGGGRQRLQESFQPEDNCG 611
Db      182 PYLEQHRG 189
OY      612 SVVPGQGG 619
Db

```

Search completed: August 27, 2003, 17:45:30
 Job time : 55 secs